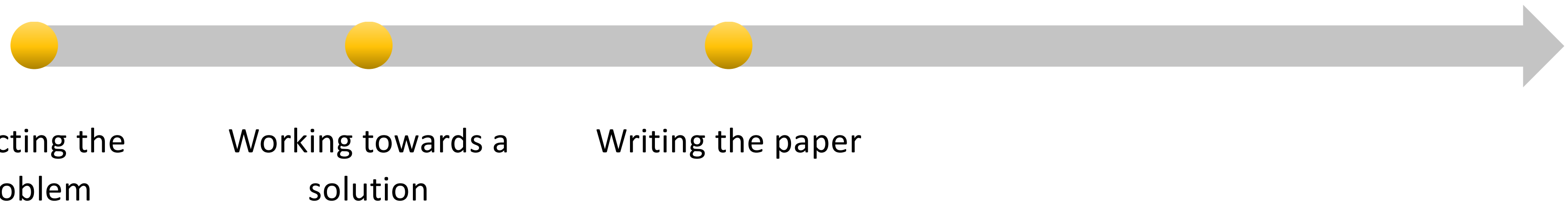
From Ideas to A* Papers:

A Behind-the-Scenes Journey not Often Discussed

Joseph K J

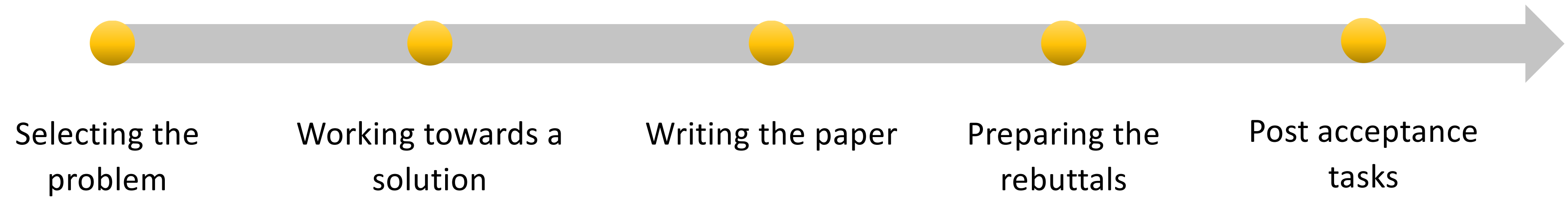
Observations might be biased towards Computer Vision and Machine Learning conferences.

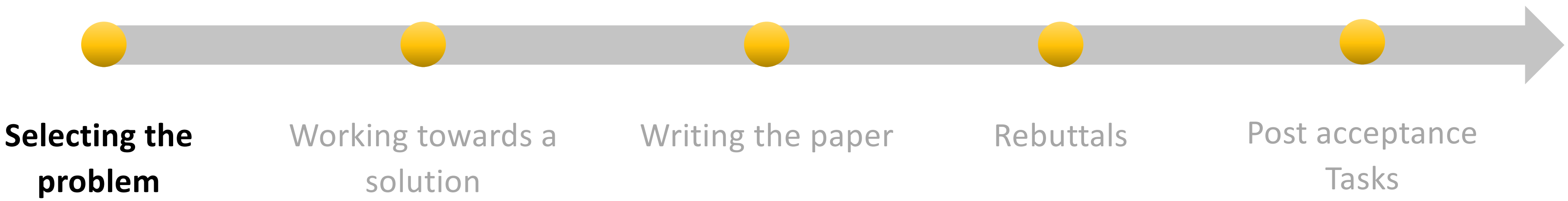
Milestones



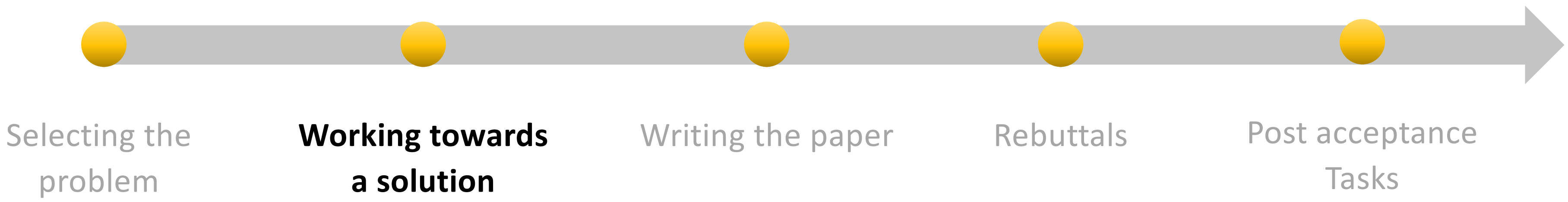
Selecting the problem
Working towards a solution
Writing the paper

Milestones


Selecting the problem
Working towards a solution
Writing the paper
Preparing the rebuttals
Post acceptance tasks



- Do something that you are really passionate about.
  - It’s a roller-coaster, and your passion is your seat-belt
- Own the problem
  - It is your PhD thesis
- Practical tips
  - Select a topic that is not saturated.
  - Is of interest to the venues that you are targeting.
  - Even early on, shoot for the best venues.
  - Find the baseline paper that you are going to build on, ensure that the code is available.



- Setup your baseline.
- Iterate between coding and reading papers.
- Be organized with your code: use GitHub.

github.com/JosephKJ/OWOD/commits/master/?before=23890f188cd1a6801c6ac0e3dacd78b8572b8c29+70

Open in app | Paused | New Chrome available

master | All users | All time

Commits on Nov 12, 2020

- iOD
JosephKJ committed on Nov 12, 2020 — 68da418

Commits on Nov 9, 2020

- Saving feature store
JosephKJ committed on Nov 9, 2020 — 81516e5

Commits on Nov 6, 2020

Weibull distribution: Inference added
JosephKJ committed on Nov 6, 2020 — ea9f7d1

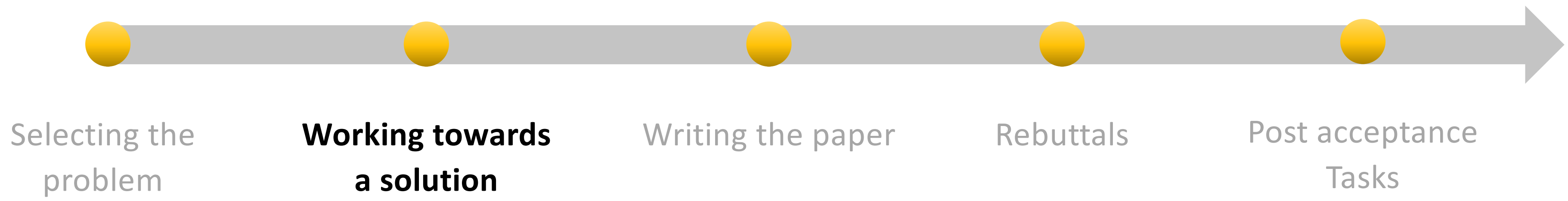
Weibull distribution
JosephKJ committed on Nov 6, 2020 — 71480aa

Commits on Nov 4, 2020

Updated the evaluation display. Ready for t1_std_frcnn.
JosephKJ committed on Nov 4, 2020 — 9321996

- Adding new imagesets
JosephKJ committed on Nov 4, 2020 — 92312cc

Commits on Nov 3, 2020



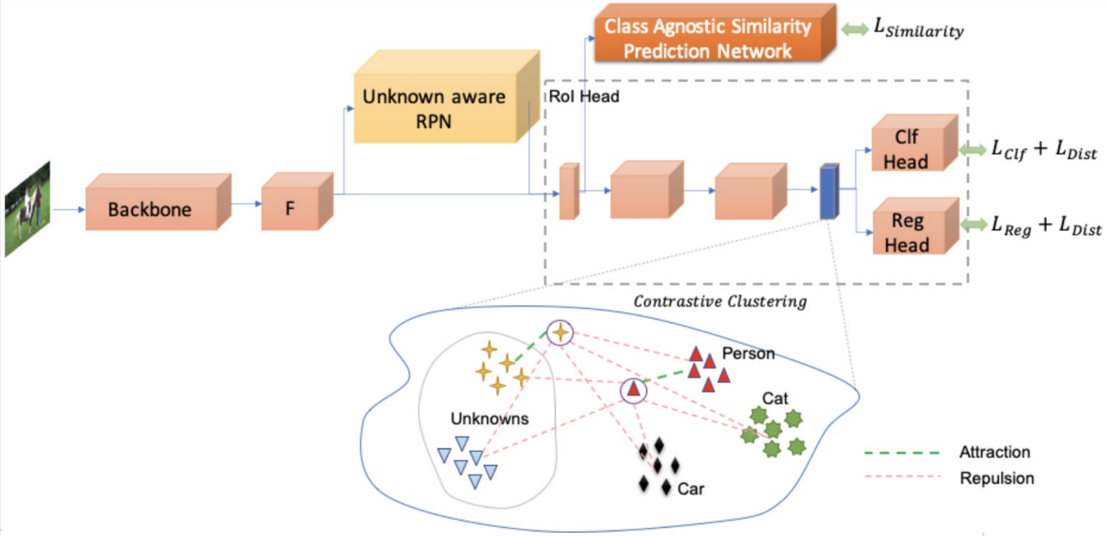
- Setup your baseline.
- Iterate between coding and reading papers.
- Be organized with your code: use GitHub.

- Stuck? First, talk to yourself.
  - Maybe a white-board
  - Write things down
  - Then discuss with others
    (don't expect anything from anyone, if you get, be grateful)
- Be organized with your work: log it.
  - Be prepared to be disturbed.

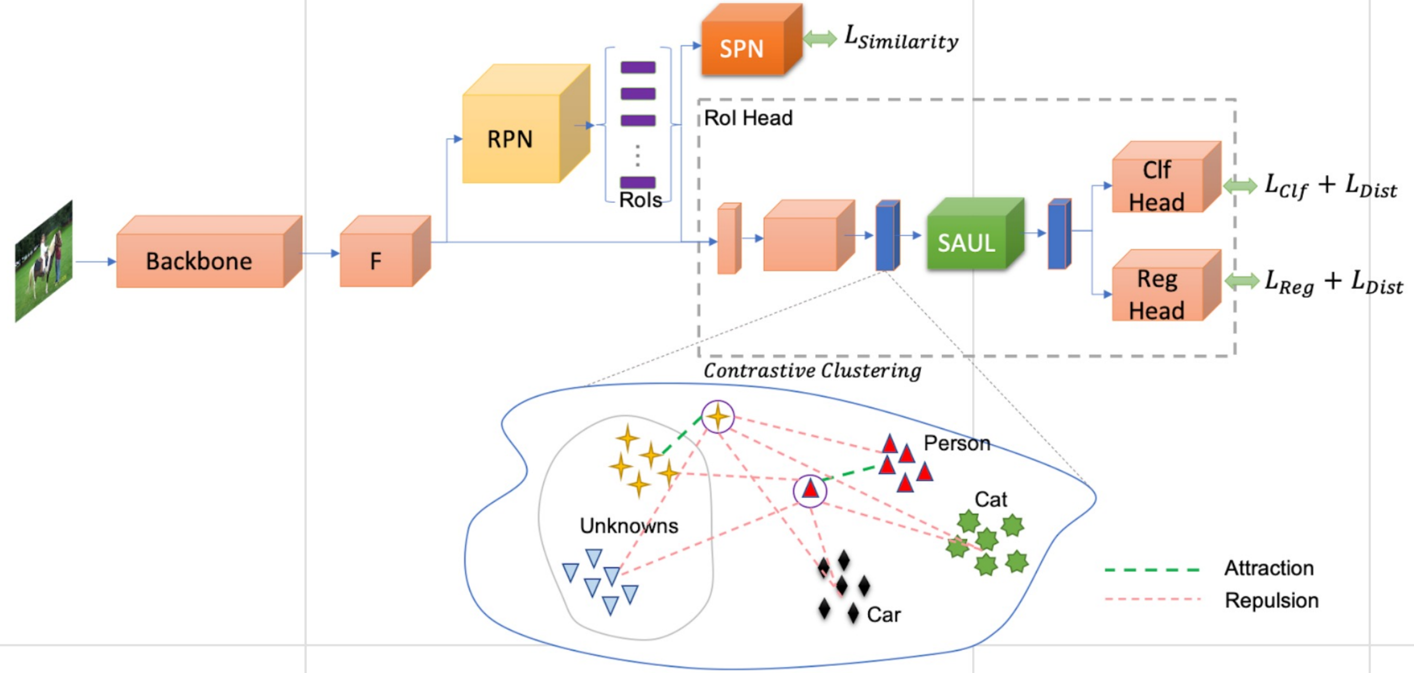
	A	B	C	D	E	F
248	**1) Mar 15 - Mar 21**					
249	**Papers**					
250	1	Continual Unsupervised Representation Learning	NeurIPS 19	Introducing "Unsupervised Continual Learning":	MNIST	
251				--- Task label and boundary is unknown. (Task aware setting)	Omniglot	
252				--- Each datapoint is not assigned label (Un-supervised Learning)		
253				A mixture of gaussian is assumed where each component is specific to each task. Each component is dynamically added.		
254				An extension to Variational Deep Embedding.		
255						
256	2	Unsupervised Continual learning and Self-Taught Associative Memory Heirarchies	LLD Workshop	They introduce a new STAM architecture that rivals CNN.		Problem definition is similar
257			ICLR 2019	Slight few-shot flavour to the problem, as during inference, few labelled examples are used.		
258						
259	3	Automatically discovering and learning new visual categories with ranking statistics	ICLR 2020	Primariliy focused on (un)supervised clustering of novel classes; a set of labelled base class images are available.		
260				Use RotNet to get a feature representation which is finetuned.		
261				Psuedo Labels are generated for the novel class images.		
262						
263						
264	**Summary of main points discussed in the meeting (Salman)**					
265	[17 Mar 20]	‣ Incremental Few-shot: See how we compare against the CVPR 2020 paper.				
266		‣ Extending the ECCV Submission				
267		→ CenterNet, RetinaNet				

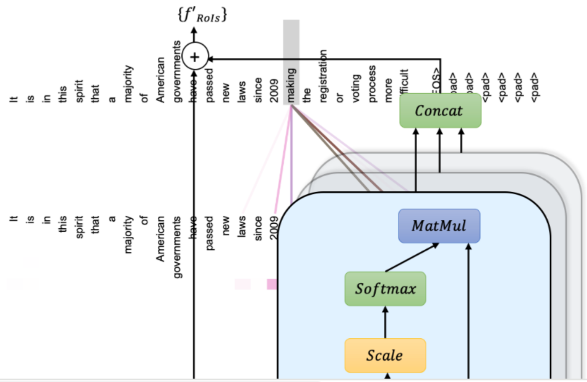
Updates | Manuscript | Timeline | Reporducing results | Supplementary Section | Items for clustering | Momentum | Margin | 1 Data

	A	B	C	D	E	F	G	H
1	Idea 1	Auto-label unkowns and learn						
2		Self attentive maps						
3								
4	Idea 2	BNN instead of classifier and regressor						
5		- Uncertainly modelling.						
6								
7	Idea 3	Expore Outlier detection methodoloies						
8								
9								
10	Idea 4	Inspiration from Active Learning?						
11								
12								
13	Idea 5	Would edge detection help for unknwon detection as its class agnostic?	https://arxiv.org/pdf/1902.10903.pdf					
14								
15								
16	Idea 6	There are these papers which says that finetuninig with exemplar set is all that is required.						
17		- GDumb, ECCV 2020	https://www.ecva.net/papers/eccv_2020/papers_ECCV/papers/123470511.pdf					
18		- Frustratingly Simple Few-Shot Object Detection, ICML 2020	https://proceedings.icml.cc/static/paper_files/icml/2020/2957-Paper.pdf					
19		Hence, why dont we just finetune to prevent Catastrophic forgetting with a small subset of data. How we select the data would be important then.						
20								
21								
22								
23								
24								

	A	B	C	D	E
1	Highlevel phases: Image -> Backbone -> Unknown Aware RPN -> Similarity Prediction Network -> Contrastive Clustering -> MixDist				
2					
3	(NB: All citations refer to the papers in the first sheet.)				
4	**Component**	**Methodology**	**Positives**	**Negatives**	**Additional ways to impro**
5	Unknown Aware RPN	Label top-k (based on objectness score) background proposals as Unknown.	Easy way to auto-label an unknown. Qualitative results are coherent.	- Hyper-parameter k, i, o	1. Some boundary detectio loss to improve the objectn score.
6		Background proposals are those with IoU < i		- Seems to be a simple hack.	2. Better Uncertainity estim methodology (similar to the latent exploration method)
7					
8					

2 Methodologies | 2 Methodology-v1 | 1 Methodology-v2 | Wilderness Results | Evaluation Metric | TODO | (old) Results | (old) Quant

	A	B	C	D	E
1	Highlevel phases (v1): Image -> Backbone -> Unknown Aware RPN -> Similarity Prediction Network -> Contrastive Clustering -> MixDist				
2	Highlevel phases (v2): Image -> Backbone -> RPN -> Similarity Prediction Network -> Contrastive Clustering -> Self Attentive Unknown Labeller -> ~~MixDist~~ Exemplar replay + Distillation				
3					
4					
5	**Component**	**Methodology**	**Discussion**	**Positives**	**Negatives**
6	RPN	The standard RPN, which produces region proposals (objectness score and BB offsets).			
7		It is class agnostic			

	A	B	C	D	E
11					
12		**Intuition is that the unknown object will be attended to by the known objects.** To be more explicit, we can assume that some of the background region proposals would be attentded more (measured in terms of the number of in-bound attention links (hyperparam in_link)). Such background proposals can be labelled as unkown.			hyperparam: in_link
13		Architecture of SAUL: 			

Vineeth N Balasubramanian
25 Sept 2020

What is the logic for this? Do we mean that these are more likely to mislead object detectors as being obhects, than a more amorphous background?

Deleted user
25 Sept 2020

Yes, and even more importantly, such unknown object might be automatically discovered.

Comments above copied from original document

2 Methodologies | 2 Methodology-v1 | 1 Methodology-v2 | Wilderness Results | Evaluation Metric | TODO | (old) Results | (old) Quant

Selecting the problem | Working towards a solution | **Writing the paper** | Rebuttals | Post acceptance Tasks

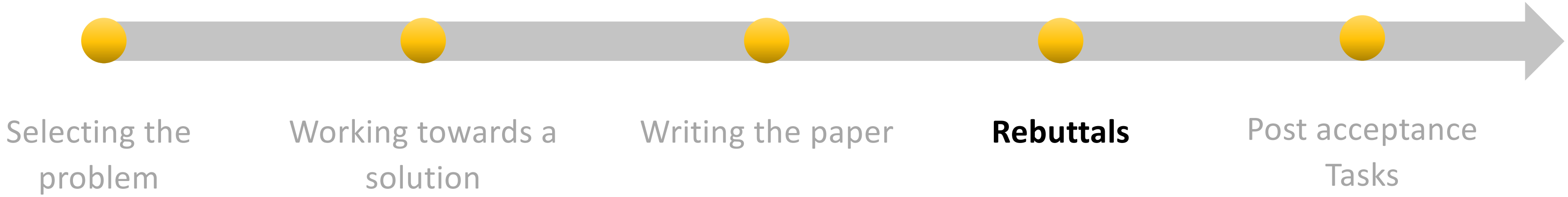
- Plan your writing: have a skeleton first.

	B	C	D	E	F	G	H	I
13	**Open World Object Detection**							
14	Related Wc	Methodology		Experiments and results				Conclusion
15		Problem Formulation	Proposed Solution	Evaluation protocol	Implementation details	Main Results	Additional results	
16			- Contrastive clustering	- Data split	8 Tesla V100 GPUs	- Ability to reduce the confusion of an unknown being labelled as a known.	- Main Ablations	
17			- RPN Thresholding based auto-labelling	- Evaluation metrics	Batch size: 8; one image per GPU	- Ability to explicitly identify an unknown	- Time taken: training and inference	
18			- Energy based unknown identification		Custom dataloaded which labels images accoding to whether it is known or unknown	- Incremental OD results	- Plot of clustering loss.	
19			- Balanced finetuning for addressing catastrophic forgetting. (Mitigating forgetting)		While training T_i, only classes in T_i are labelled, all the others are ignored.		- TSNE of latent space	
20					How the extra class of backgrounds are added		- Qualitative results	
21			- A unified solution to OWOD, which is a symphony of		Handling incremental class growth.		- Memory Size	
22							- Confusion matrix?	
23	Open woRld dEtectOr						- Contribution of mixed unknown errors to the whole error.	
24							- Modelling the uncertainity: weibull distribution.	
25								
26								
27	ons							

	B	C	D	E	F	G	H	I	J	K	L
23	**Method**					Scorer: Measuring Slide's Quality					
24	☑		Why RL to improve slide generation?								
25	☑			- Dataset for SFT is impractical to collect.							
26	☑		How to adapt RL with noisy / imperfect rewards								
27	☐										
28	☐	**Coactive Learning a Slide Stylizer**									
29	☑		**Preliminary**								
30	☑			Takeway: in expectation, if we have an improved version of the slide, it will help.							
31	☐			We should introduce π_*							
32	☐		**Slide Stylizer**								
33	☐			We plan to learn a **single module compising of planner and a coder** that can take in components, current rendition and generate a stylized							
34	☑			~~This single module is what we learn with coactive learning.~~							
35	☑			Creating π_* similar to Tucker et al. needs human annorated data, but this is infeasible for us. Instead, we propose a way to sample and strictl							
36	☑			**Building a robust π_***							
37	☐				How to generate improved feedback (π_*)?						
38	☑					Qwen-VL - ZS					
39	☑					Qwen-VL - FT or	These models are trained only to make slight improvement, which greatly simplifies that ta				
40	☐				How do we measure improvement?						
41	☐					The scorer					
42	☐					Show that it is good.					
43	☑			**Building single module compising of planner and a coder**							
44	☑				Data generation pipeline						
45	☑				SFT						
46	☐				DPO						

CVPR Timeline | **Skeleton** | AutoPresent-Slidev | AutoPresent Evaluation | Pairwise-he | Sheet19 | Todo | SFT Results | DPO Re

- Plan your writing: have a skeleton first.
- First draw the figures and write any algorithms (if you may have).
- Certain parts of the paper can start early
  - Related works: during your initial literature review phase
  - Intro: have pointers early on.
- Plan to finish at-least one week before the deadline.
  - Iterate through it, to make it perfect.

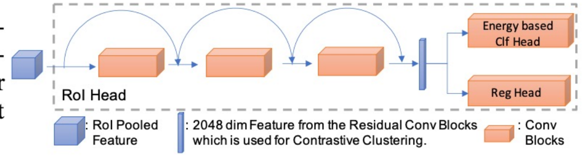


- Yes, it is possible to turn-around the papers with a clear rebuttal

We thank the reviewers for the positive feedback: *novel approach and problem setting (**R1**, **R3**), core ideas are interesting and worth explored (**R2**), addressed problem is very hard and very important for AI (**R2**), well-written (**R1, R2, R3**), literature survey is up to date and clearly explained (**R1, R3**), evaluation is well conducted (**R2**), extensive experiments and ablations to showcase different aspects (**R3**).* Our source code and models will be publicly released.

Reviewer 1 (R1) **Name (OREO):** Thanks. We will rename our method. **Relation with embedding learning (EL):** Different from EL works, our approach performs clustering by using contrastive constraints on *dynamic prototypes from each class* (Eq 1) instead of single class instances. This is important for our OWOD problem since same class instances are desired to be mapped together while they must be far from *all other class instances* simultaneously. To our knowledge, we are the first to introduce a multi-way class-level contrastive formulation for the challenging OWOD problem. Further, contrastive constraints in most existing EL methods (*e.g.*, He *et al.*, MoCo, CVPR'20) have higher complexity and longer training times due to *instance level*

Figure A1. RoI head architecture, showing 2048-dim feature vector used for contrastive clustering. Best view zoomed in.



fold: 1) it enables the model to cluster unknowns separately from known instances, thus boosting unknown identification; 2) it ensures instances of each class are well-separated from other classes, alleviating the forgetting issue (L266). The 2048-dim feature vector that comes out from residual blocks of RoI head (Fig A1) is contrastively clustered (L592). The contrastive loss is added to the Faster R-CNN loss and the entire network is trained end-to-end (L371). We will clarify in revision. **Role of Energy:** The intrinsic capability of EBMs [23] to assign low energy values to in-distribution data and vice-versa motivates us to use an energy measure to characterize whether a sample is from an unknown class. **Clarification regarding Unknown Identification:** An unknown is *not identified by thresholding* in our case, but by computing the likelihood of the data point w.r.t a learned Weibull distribution, which is fit only once at inference and takes just 1.78 sec. **Comparison with**

COCO) as 'unknown', we need to re-annotate all images, which is laborious, as it takes ~ 20k worker hours just for point annotation on COCO [27]. Auto-labeling region proposals with our method is an effort to address this problem (Sec 4.2), but is not a substitute for ground-truth since pseudo labels are noisy. However, we show that our method can use these noisy pseudo labels to aid latent space separation via contrastive clustering, which helps better classification by the energy-based unknown identifier. **On Keeping all Data:** The storage and compute expense will grow unbounded if we are to store all data and retrain a model from scratch each time. We follow standard incremental learning settings [1,4,5,20,21,24,32,43,44,52,57] in this regard, which work without storing all data. We will clarify this in revision. **Difference from Open Set/World:** Fig 1 shows how our OWOD problem differs to Open Set/World. We would like to borrow R2's comment that OWOD setting is '*very hard and very important for AI*'. **Additional Baseline:** As suggested, we compare our method with [2, 3] in Tab A1 and will include these results in the revision.

Method	WI (↓)	A-OSE (↓)	mAP-KU (↑)	mAP-K (↑)
Objectness[2]	0.089	14892	12.15	27.42
Deepbox[3]	0.069	13396	34.18	53.23
OWOD	**0.022**	**8234**	**56.34**	**81.31**

Table A1. Baseline comparison with [2, 3]. mAP-K: Only Known. mAP-KU: mAP on Known & Unknown. Best view zoomed in.

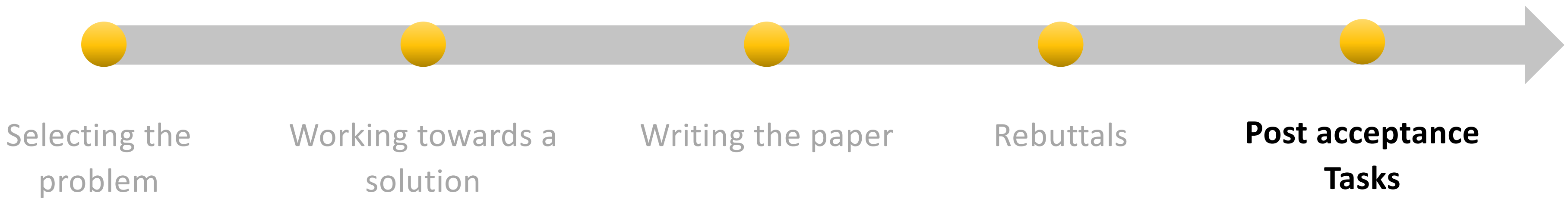
Reviewer 2 (R2) More Details on Contrastive Clustering: The motivation for using contrastive clustering is two-

Feature Compatibility: Our balanced finetuning helps improve feature compatibility. Thank you for sharing the idea of fixing the classifier. We will explore this promising direction as a next step and cite Shen *et al.*, Pernici *et al.*

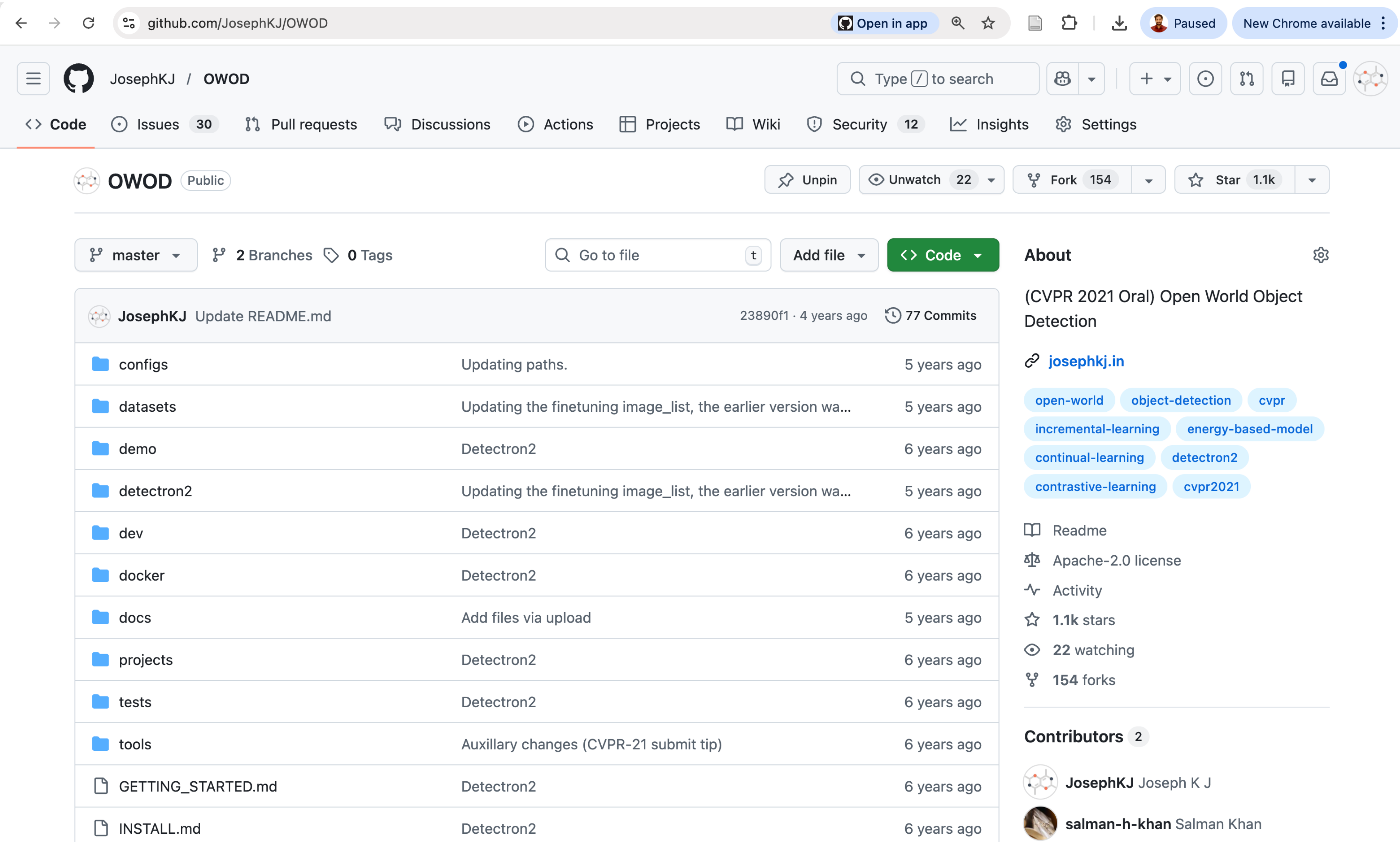
Reviewer 3 (R3) Failure Cases: Occlusions and crowding of objects are cases where our method tends to get confused (*storage*, *walkman* and *bag* not detected as *unknown* in Figs. S5, S7). Difficult viewpoints (such as backside) also lead to some misclassifications (*giraffe*→*horse* in Figs. 4, S6). We have also noticed that detecting small *unknown* objects co-occurring with larger known objects is hard. We thank R3 and will include detailed failure analysis in revision. **Recall Threshold:** A high recall (0.8) was set to evaluate in a challenging scenario. At 0.5, we observe a slight gain in WI to 0.019. The change is not drastic since WI computation involves ratio of the precisions at a given recall. **Temperature:** We fixed T=1 (L121 in Suppl.). Softening the energies a bit more (T=2) gives slight improvement in unknown detection (A-OSE and WI), however increasing it further hurts as evident from adjacent table.

T	WI(↓)	A-OSE(↓)	mAP(↑)
1	0.0219	8234	**56.34**
2	**0.0214**	**8057**	55.68
3	0.0411	11266	55.51
5	0.0836	12063	56.25
10	0.0835	12064	56.31

Energy Motivation: Kindly see R2 (Role of Energy). **Other Issues:** We thank R3 and will cite and discuss recent detectors. Supplementary will also be referred adequately.



- Make a webpage, YouTube video etc.
- Made your code and models public.


github.com/JosephKJ/OWOD
Open in app
Paused
New Chrome available
JosephKJ / OWOD
Type / to search
Code
Issues 30
Pull requests
Discussions
Actions
Projects
Wiki
Security 12
Insights
Settings
OWOD
Public
Unpin
Unwatch 22
Fork 154
Star 1.1k
master
2 Branches
0 Tags
Go to file
t
Add file
Code
JosephKJ Update README.md
23890f1 · 4 years ago
77 Commits
configs Updating paths. 5 years ago
datasets Updating the finetuning image_list, the earlier version wa... 5 years ago
demo Detectron2 6 years ago
detectron2 Updating the finetuning image_list, the earlier version wa... 5 years ago
dev Detectron2 6 years ago
docker Detectron2 6 years ago
docs Add files via upload 5 years ago
projects Detectron2 6 years ago
tests Detectron2 6 years ago
tools Auxillary changes (CVPR-21 submit tip) 6 years ago
GETTING_STARTED.md Detectron2 6 years ago
INSTALL.md Detectron2 6 years ago
About
(CVPR 2021 Oral) Open World Object Detection
josephkj.in
open-world
object-detection
cvpr
incremental-learning
energy-based-model
continual-learning
detectron2
contrastive-learning
cvpr2021
Readme
Apache-2.0 license
Activity
1.1k stars
22 watching
154 forks
Contributors 2
JosephKJ Joseph K J
salman-h-khan Salman Khan

TITLE	CITED BY	YEAR
Towards open world object detection KJ Joseph, S Khan, FS Khan, VN Balasubramanian Proceedings of the IEEE/CVF conference on computer vision and pattern …	789	2021

github.com/JosephKJ/OWOD/issues?q=is%3Aissue%20state%3Aclosed&page=2 Open in app Paused New Chrome available

Confusion about output of RPN block
#84 · by giangnt071098 was closed on May 12, 2022 — 1

KeyError: 'Non-existent config key: OWOD'
#83 · by chengtao-lv was closed on Jan 19, 2022 — 2

Unable to determine the device handle for GPU 0000:02:00.0: GPU is lost. Reboot the system to recover this GPU
#81 · by JohnWuzh was closed on May 12, 2022 — 1

ValueError: min() arg is an empty sequence & t1_clustering_with_save
#80 · by CtCCtV was closed on May 12, 2022 — 5

Not able to detect unknowns
#79 · by Jiyang-Zheng was closed on Dec 24, 2021 — 3

Where is the code of "4.3. Energy based unknown identifier"
#78 · by lsqxxx was closed on Dec 7, 2021 — 1

I cannot reproduce the result either. Please help, thank you very much!
#77 · by YujunLiao was closed on May 12, 2022 — 2

Confusion about the t2_train/t2_ft procedure
#76 · by YujunLiao was closed on May 12, 2022 — 1

万事俱备只欠东风
#75 · by Allenstin was closed on Dec 4, 2021 — 17

This project,how i train my own data and test?
#74 · by guyue1994 was closed on Dec 4, 2021 — 2

How t1_train.txt is generated.
#73 · by lxqiaoyixuan was closed on Dec 4, 2021 — 2

question about energy .png
#72 · by PowderYu was closed on May 12, 2022 — 1

github.com/JosephKJ/OWOD/issues/40

Open in app | Paused | New Chrome available

Code | Issues 30 | Pull requests | Discussions | Actions | Projects | Wiki | Security 12 | Insights | Settings

Same START_ITER values in t*_train.yam and t*_ft.yaml #40

Edit | New issue

Closed

mmaaz60 opened on May 15, 2021

Hi,

I noticed that the value of `START_ITER` is the same in all corresponding `train` and `ft` configs. For example, it is `START_ITER: 18000` in both `t2_train.yaml` and `t2_ft.yaml`. As per the script run.sh, the finetuning (`t*_ft.yaml`) is using the weights generated from the corresponding `t*_train.yaml` as starting point, so the `START_ITER` shouldn't be the same in both files. Please guide.

Thanks

Create sub-issue

JosephKJ on May 15, 2021 — Owner

Hi Maaz,

The START_ITER that you are referring to is the iteration from which we start the clustering process. Kindly note that we dont do clustering while finetuning. Kindly see this flag.

I am afraid that you are confusing this with MAX_ITER, which controls the number of iterations that the model is trained for.

Kindly let me know whether this clarifies your doubt.

Thanks,
Joseph

Assignees
No one - Assign yourself

Labels
No labels

Projects
No projects

Milestone
No milestone

Relationships
None yet

Development
Code with agent mode

Create a branch for this issue or link a pull request.

Notifications — Customize

Unsubscribe

github.com/JosephKJ/OWOD/issues/35

Code | Issues 30 | Pull requests | Discussions | Actions | Projects | Wiki | Security 12 | Insights | Settings

Reason for the slow response last month #35

Closed

JosephKJ opened on May 8, 2021 — Owner

Hi all,

My apologies for the slow response during April 2021. I was seriously hospitalized with COVID pneumonia. I lost 20 % of my lungs to the virus. I am improving now. Please find the discharge summary attached along, if that helps you to understand better.

discharge_summary.pdf

Thanks,
Joseph

Create sub-issue

JosephKJ mentioned this on May 8, 2021

[resolved] Can anyone reproduce the results? (my results attached) #26

JosephKJ closed this as completed on May 11, 2021

qianyizhang on Sep 17, 2021

stay strong!

Assignees
No one - Assign yourself

Labels
No labels

Projects
No projects

Milestone
No milestone

Relationships
None yet

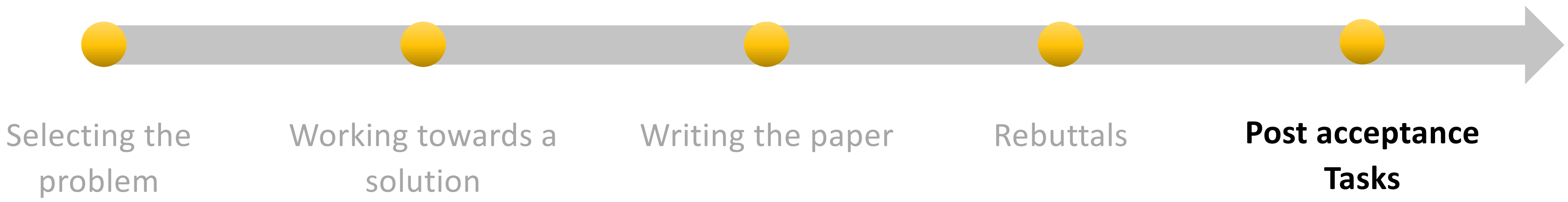
Development
Code with agent mode

Create a branch for this issue or link a pull request.

Notifications Customize

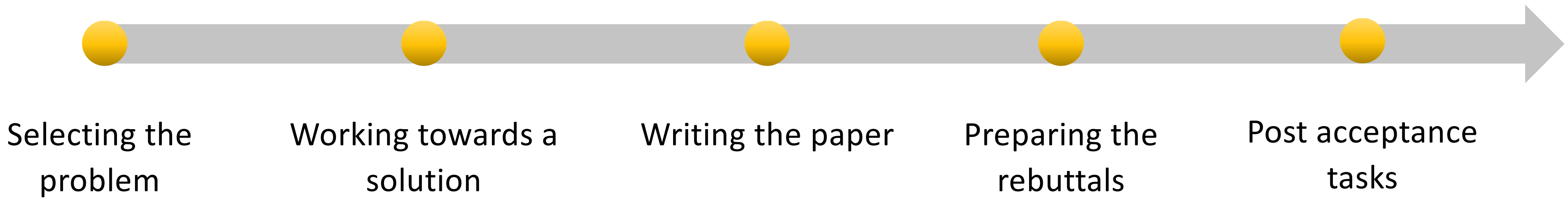
Unsubscribe

You're receiving notifications because you're



- Make a webpage, YouTube video etc.
- Made your code and models public.
- Prepare for your talk and presentation well: giving a 5 mins talk is harder than giving a 30 mins talk!
- Plan your travel!

Milestones


Selecting the problem
Working towards a solution
Writing the paper
Preparing the rebuttals
Post acceptance tasks

Failures are blessing-in-disguise

The whole system is noisy, please don't take it personally. Improve your work and move on.

Don’t give up easily

If you truly love what you are doing, and own it, you will not let go of it.

Vineeth N Balasubramanian

భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Salman Khan

Fahad Khan

جامعة محمد بن زايد
للذكاء الاصطناعي
MOHAMED BIN ZAYED UNIVERSITY
OF ARTIFICIAL INTELLIGENCE

PhD days are the best phase of your life, enjoy every bit of it!

Thank you!