Grading Policy and Requirements of Course Project

Yue HU, Chenxin XU

6/13/22

Grade component

Project takes 36% of your final grade.

Main component	Ratio	Requirement
Team-up report	4%	Finished
Proposal	4%	• Decide the challenge to participate in. • Draft the system design. • Check and prepare the dataset.
Progress report	4%	• Finish 50% of the work. • Implement the major part of your method.
Presentation	10%	Deliver your work in the form of video (5min).
Final report	14%	Complete experiments and final report.

01 Competition Track

Detailed grading policy

Proposal (4%)

- Challenge description
  - Describe the basic information about the challenge.
    Size of the dataset, the baselines, the organizer, the number of participated teams, etc.
  - Illustrate the reason why you choose this competition.
- Detailed plan
  - Propose the timeline to accomplish the project.
  - Specify the details about the dataset, baselines, and evaluation metrics.
  - State the method you plan to use.
- Roles and responsibilities of each team member in this project

How to choose your challenge?

- **Data:** open source; a reliable quality; more than 5000 samples.
- **Task:** meaningful with useful applications.
- **Number of participated group:**
  - At least 20 teams to participate.
    A larger competition scale is encouraged (like > 100 teams).
- **Baselines:**
  - At least 20 valid baselines(methods) on the leaderboard.
    Causal submissions (for example, a submission with <50% accuracy to a classification task) are not considered as valid baselines.

Please discuss with the instructors before you decide your challenge!

Detailed grading policy

Progress report (4%)

- Progress
  - Describe the overall progress of your project.
    You are expected to finish 50% of your project at this point.
  - Report the difficulties you met, and your plan to handle.
- Understanding
  - Present your understanding of the baselines and evaluation metrics.
  - Demonstrate the implementation details of your method and explain why.
- Writing: grammar, logic, clarity
- Contribution of each team member in the project

Detailed grading policy

Final report (14%) the critical level of each factor is ranked in a decreasing manner.

1. Performance of your methods (7%)
2. Improvement over the baseline method
3. Evaluation and analysis
4. Writing
5. Demonstrating strong understanding

Detailed grading policy

Final report (14%) - Performance of your methods (7%)

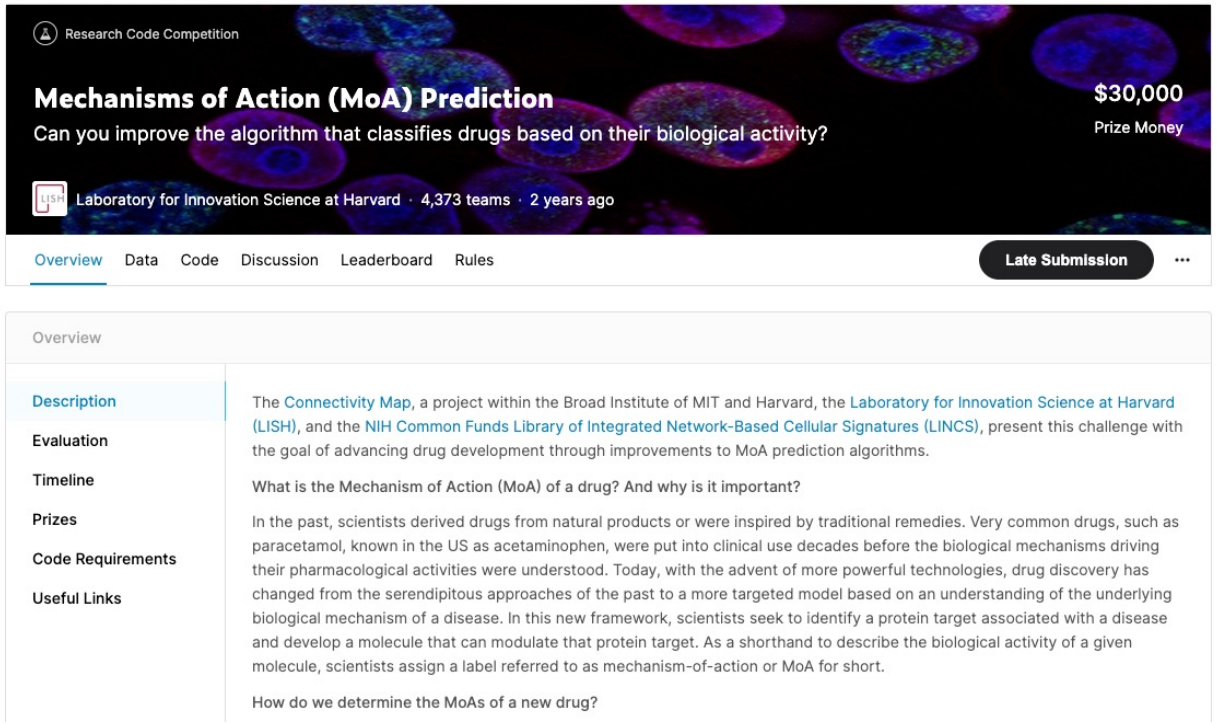
case by case

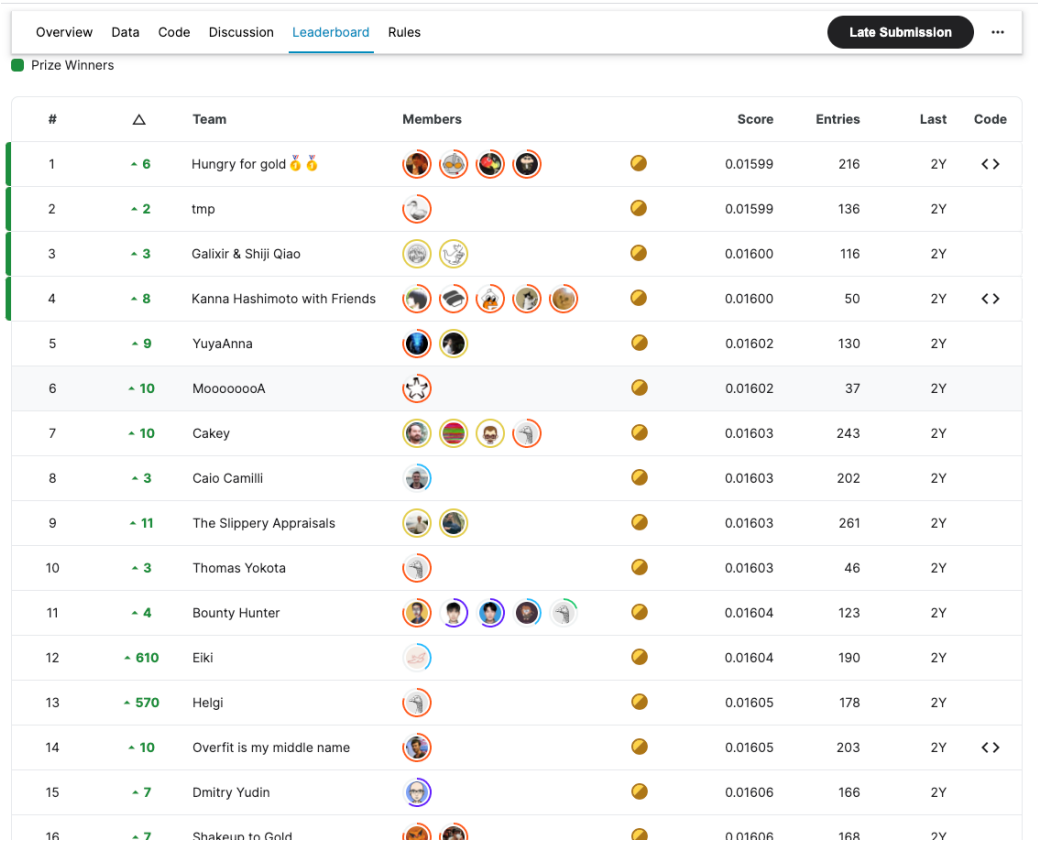
Competition quality	Too simple/Unreliable	Fair	Challenging
Rank 1~5%	6%	7%	7%
Rank 6%~20%	5%	6%	7%
Rank 21%~40%	4%	5%	6%
Rank 41%~60%	3%	4%	5%
Rank 61%~100%	2%	3%	4%
Invalid submission	0%	0%	0%

Note: directly using an existing open-source baseline will lead to a very low score although the performance is good

Task samples

1. Mechanisms of Action (MoA) Prediction





>4000 submissions

Task samples

2. NuScenes 3D detection challenge

Rank	Participant Team	mAP (↑)	mATE (↑)	mASE (↑)	mAOE (↑)	mAVE (↑)	mAAE (↑)	NDS (↑)	Submitted at
1	HANLAB (BEVFusion-e)	0.75	0.24	0.23	0.32	0.22	0.13	0.76	8 day(s) ago
2	RacoonPower	0.74	0.23	0.23	0.29	0.24	0.13	0.76	72 day(s) ago
3	UNISOC_MMLab (FusionVPE)	0.73	0.24	0.23	0.28	0.24	0.13	0.75	21 week(s) ago
4	DAA-MRC (DAA)	0.73	0.23	0.23	0.28	0.26	0.13	0.75	86 day(s) ago
5	Robotics and Autonomous Driving Lab (CenterPoint-Fusion)	0.72	0.24	0.23	0.32	0.21	0.13	0.75	54 week(s) ago
6	JDXAD (PAI3D)	0.71	0.24	0.23	0.31	0.23	0.13	0.74	23 week(s) ago
7	FocalsConv (FocalSparseCNN)	0.70	0.24	0.23	0.31	0.24	0.13	0.74	30 week(s) ago
8	ADS-FUSION	0.69	0.25	0.23	0.31	0.21	0.13	0.73	9 day(s) ago
9	RLV-Rush (RLVNet)	0.69	0.25	0.23	0.32	0.21	0.13	0.73	14 week(s) ago
10	LargeKernel (Large-kernel 3D CNN)	0.69	0.24	0.23	0.31	0.24	0.13	0.73	12 day(s) ago
11	DAA AVP	0.70	0.24	0.23	0.38	0.25	0.12	0.73	45 week(s) ago
12	PCIE	0.70	0.25	0.23	0.36	0.24	0.15	0.73	79 day(s) ago
13	Voxel (VFF)	0.68	0.25	0.23	0.32	0.26	0.13	0.72	30 week(s) ago
14	SJTU-VISION	0.68	0.26	0.24	0.34	0.21	0.12	0.72	88 day(s) ago
15	Alibaba_ADLab (GTFS)	0.70	0.26	0.24	0.36	0.27	0.13	0.72	72 day(s) ago
16	BEVfusion (BEVFusion)	0.69	0.26	0.25	0.37	0.27	0.13	0.72	26 day(s) ago
17	CrossFusion (DCAN)	0.67	0.25	0.24	0.31	0.28	0.12	0.72	47 week(s) ago
18	HKUST (TransFusion)	0.69	0.26	0.24	0.36	0.29	0.13	0.72	33 week(s) ago

186 submissions

Task samples

3. GQA Real-World Visual Reasoning Challenge

GQA Real-World Visual Reasoning Challenge

★ 0

Organized by: Stanford
Starts on: Feb 9, 2017, 4:00:00 AM
Ends on: Mar 2, 2099, 4:00:00 AM

Overview | Evaluation | Phases | Participate | Leaderboard | Discuss

Challenge Overview

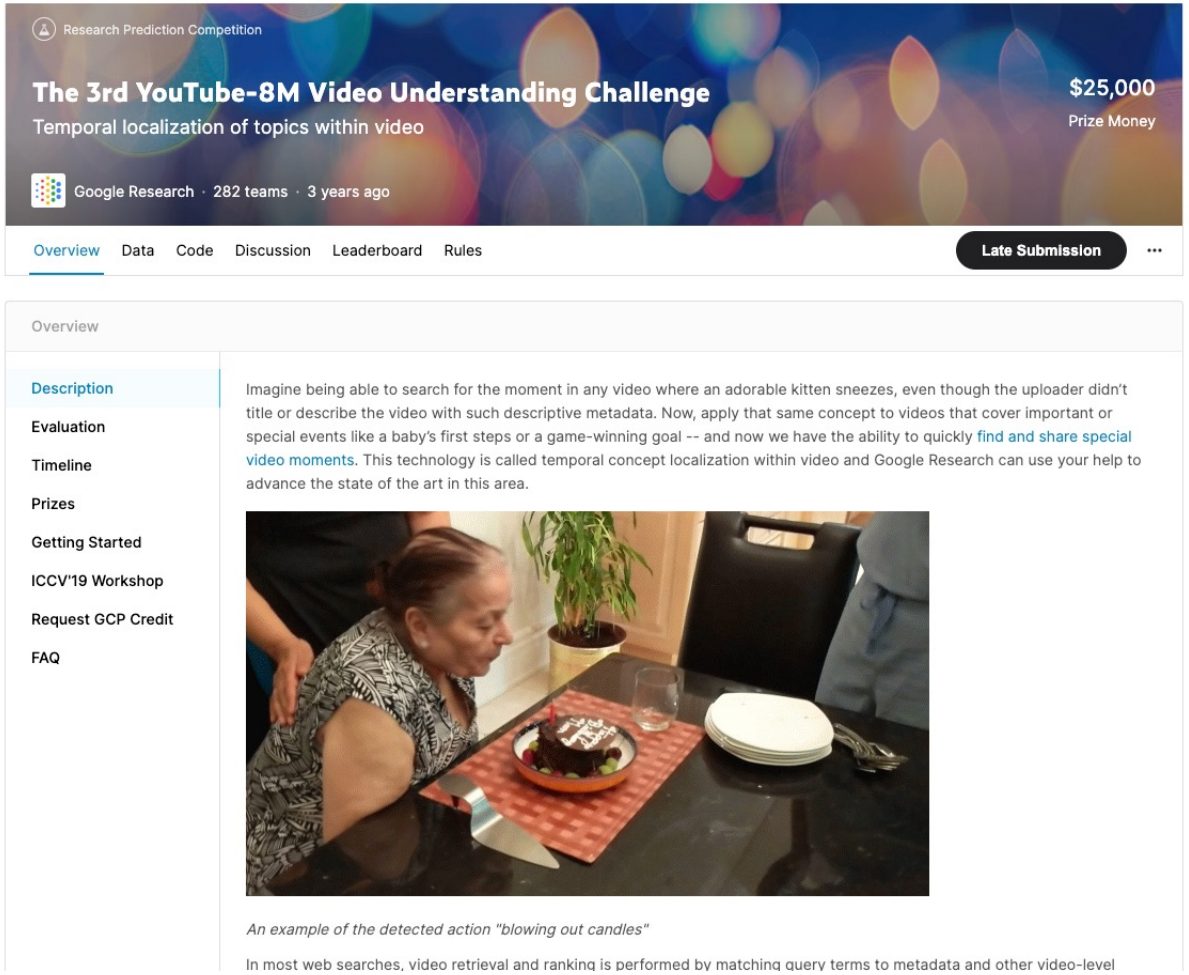
GQA is a new dataset for visual reasoning and compositional question answering on real world images. The dataset consists of 20M questions about various day-to-day images. Each image is associated with a scene graph of the image's object, attributes and relations. Each question is associated with a structured representation of its semantics, a functional program that specifies the reasoning steps have to be taken to answer it.

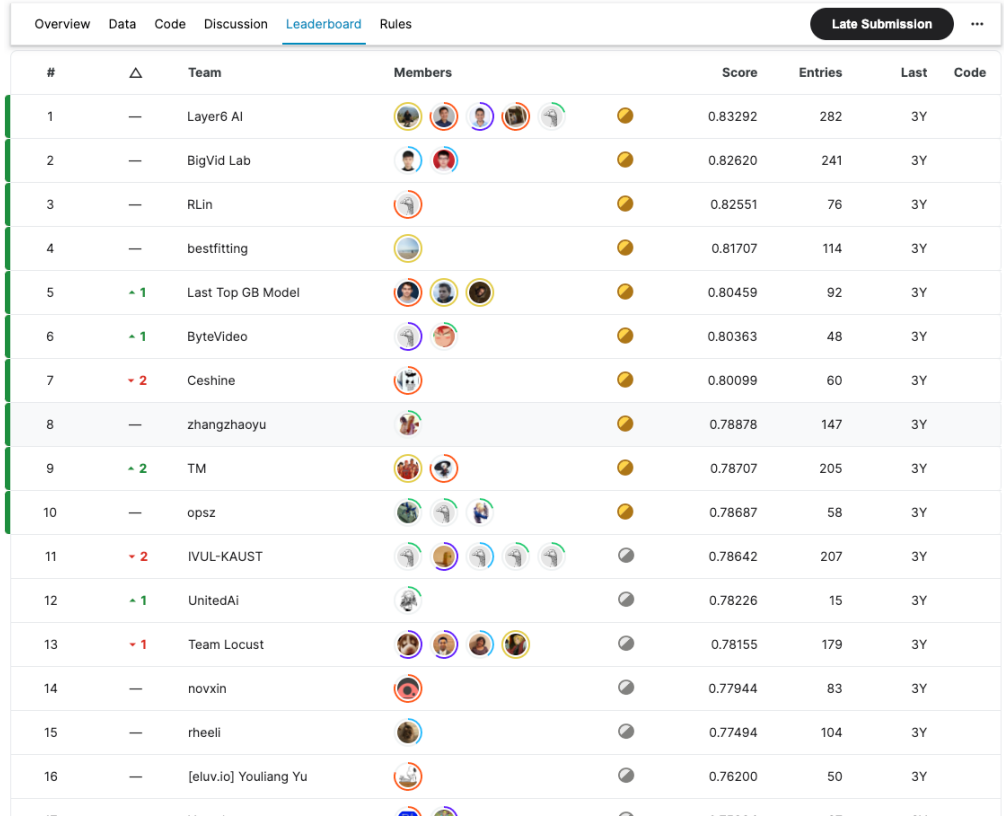
Rank	Participant Team	Binary (↑)	Open (↑)	Consistency (↑)	Plausibility (↑)	Validity (↑)	Distribution (↑)	Accuracy (↑)	Submitted at
1	Human Performance (human)	91.20	87.40	98.40	97.20	98.90	0.00	89.30	3 year(s) ago
2	DREAM+Unicoder-VL (MSRA)	84.46	68.60	91.47	83.75	96.42	3.68	76.04	2 year(s) ago
3	TRRNet (Ensemble)	82.12	66.89	89.00	83.58	96.76	1.29	74.03	2 year(s) ago
4	MIL-nbgao	80.80	67.64	91.76	83.90	96.73	1.70	73.81	91 week(s) ago
5	Kakao Brain	79.68	67.73	77.02	83.70	96.36	2.46	73.33	3 year(s) ago
6	AIOZ (Coarse-to-Fine Reasoning, Sing)	81.16	64.19	90.96	84.81	96.77	2.39	72.14	2 year(s) ago
7	270	77.50	63.82	86.94	83.77	96.65	1.49	70.23	3 year(s) ago
8	NSM ensemble (updated)	80.45	56.16	93.83	84.16	96.53	2.78	67.55	2 year(s) ago
9	liuyuhang (VinVL-DPT)	82.63	49.29	94.37	84.91	96.64	5.11	64.92	23 week(s) ago
10	VinVL (Single Model)	82.63	48.77	94.35	84.98	96.62	4.72	64.65	79 week(s) ago
11	TRRNet (Single)	77.91	50.22	89.84	85.15	96.47	5.25	63.20	2 year(s) ago
12	NSM single (updated)	78.94	49.25	93.25	84.28	96.41	3.71	63.17	2 year(s) ago
13	LXMERT (LXR955, Ensemble)	79.79	47.64	93.10	85.21	96.36	6.42	62.71	3 year(s) ago
14	MDETR (MDETR)	80.91	46.15	93.95	84.15	96.33	5.36	62.45	60 week(s) ago

>100 submissions

Task samples

4. The 3rd YouTube-8M Video Understanding Challenge

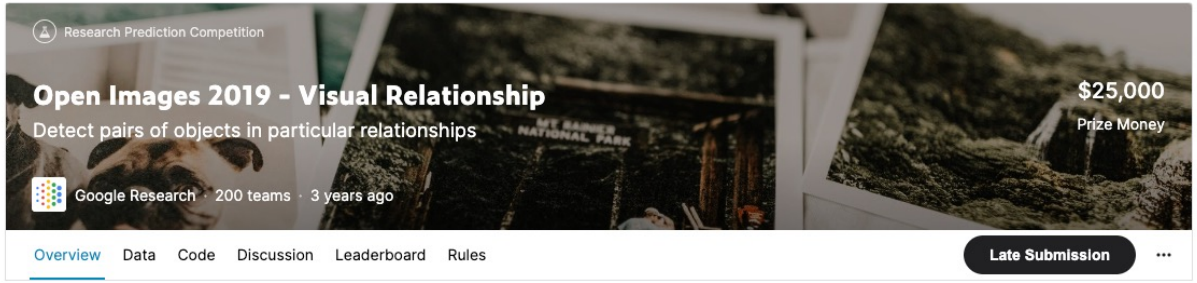




>200 submissions

Task samples

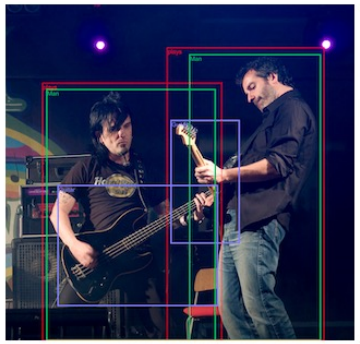
5. Open Images 2019 - Visual Relationship



Visual Relationship Track

In this track of the Challenge, you are asked to detect pairs of objects and the relationships that connect them.

The training set contains 329 relationship triplets with 375k training samples. These include both human-object relationships (e.g. "woman playing guitar", "man holding microphone"), object-object relationships (e.g. "beer on table", "dog inside car"), and also considers object-attribute relationships (e.g."handbag is made of leather" and "bench is wooden").



#	△	Team	Members	Score	Entries	Last	Code
1	—	Layer6 AI		0.40801	228	3Y	
2	—	tito		0.38818	38	3Y	
3	—	Very Random team		0.37853	162	3Y	
4	▲ 1	[ods.ai] n01z3		0.36597	20	3Y	
5	▼ 1	Ode to the Goose		0.34779	113	3Y	
6	▲ 2	X5 (⌐■_■)		0.33364	10	3Y	
7	▲ 4	JamesLearnsToMine		0.33190	71	3Y	
8	▲ 4	Ji Zhang		0.33158	1	3Y	
9	▼ 3	chicm	+4	0.32512	227	3Y	
10	▲ 3	bbbb		0.32226	16	3Y	
11	▼ 2	Ern		0.31681	94	3Y	
12	▼ 2	zgkk&sdjkn		0.30982	45	3Y	
13	▼ 6	Appian		0.30822	84	3Y	
14	▲ 5	TTX		0.26615	4	3Y	
15	▲ 3	Ivan Meleshko		0.24706	25	3Y	
16	—	winards		0.23700	10	3Y	

>200 submissions

02 Research Track

Detailed grading policy

Proposal (4%)

- Background
  - Describe the research problem; define the terminologies and concepts clearly.
  - Explain your motivation to study this problem.
- Literature review
  - List 3~5 related works, and summarize their innovation and limitations.
  - Reveal the connection of your work to the previous works.
- Research plan
  - Propose the timeline to accomplish the project.
  - Specify the details about the dataset, baselines, and evaluation metrics.
  - State the method you plan to use.
- Roles and responsibilities of every team member

Detailed grading policy

Progress report (4%)

- Progress
  - Describe the overall progress of your project.
    You are expected to finish 50% of your project at this point.
  - Report the difficulties you met, and your plan to handle.
- Understanding
  - Present your understanding of the baselines and evaluation metrics.
  - Explain your method and principle.
- **Writing:** logic, clarity, grammar
- **Contribution of each team member in the project**

Detailed grading policy

Final report (14%) the critical level of each factor is ranked in a decreasing manner.

1. Originality
   - Is this a new and appealing setting? Does this idea provide new perspective to your picked research topic? Does this idea bridge the gap or fill the blank in your picked research topic?
2. Demonstrating strong understanding
   - What is the key issue in your picked research topic? what is the weakness of previous works? what problem this paper try to tackle? And why this design could get the results as you expected?
3. Writing: concise and clear
4. Contribution of each team member in the project
5. Performance of your methods
   - Does your design outperform previous SOTA?
6. Evaluation
   - Thoroughness of your evaluation. Could the experimental results support what you claimed?

Report samples: https://openaccess.thecvf.com/CVPR2022?day=all

Detailed grading policy

Grading sample

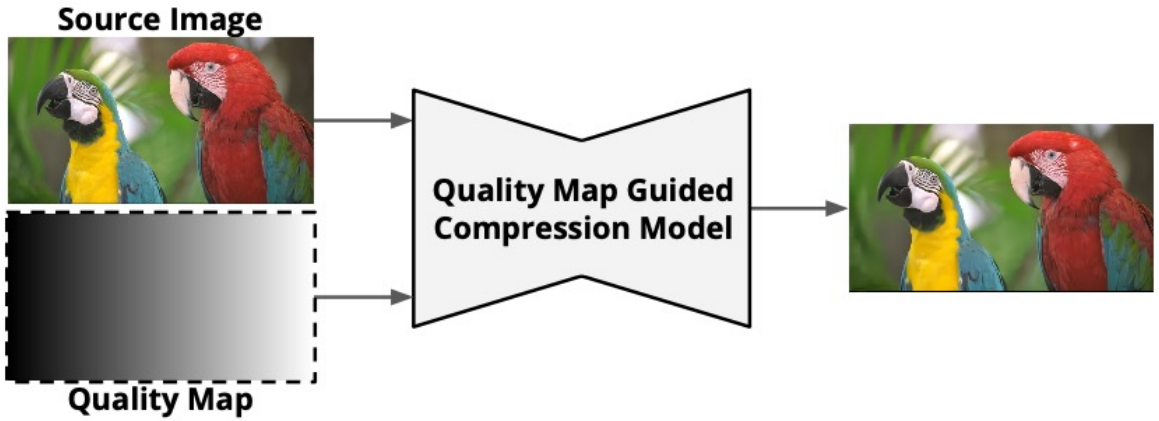
1. Re-implement one of the previous SOTAs and clearly describe the method;
    - Implement from scratch **(grade: about 60)**
    - Implement from a well-organized codebase **(grade: about 40)**
2. Conduct more detailed ablations based on the re-implemented method in i), and provide comprehensive logical analysis; **(grade: about 80)**
3. Point out one weakness of previous SOTAs, propose a novel solution and evaluate with comprehensive experiments. **(grade: about 100)**

Topic samples

1. Image compression

Task definition

Image compression is an application of data compression for digital images to lower their storage and/or transmission requirements.



References

Zhu, Xiaosu, et al. "Unified Multivariate Gaussian Mixture for Efficient Neural Image Compression." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022.

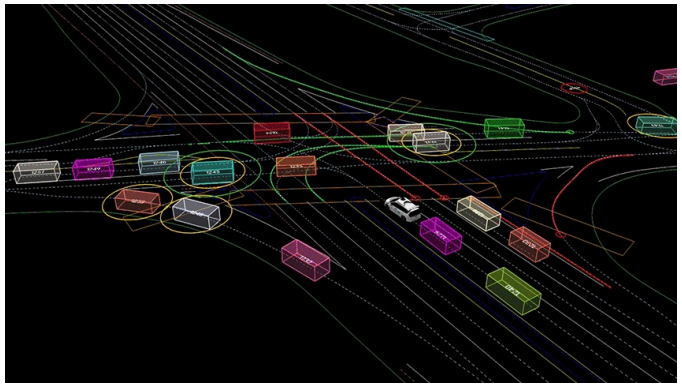
Qian, Yichen, et al. "Entroformer: A Transformer-based Entropy Model for Learned Image Compression." *arXiv preprint arXiv:2202.05492* (2022).

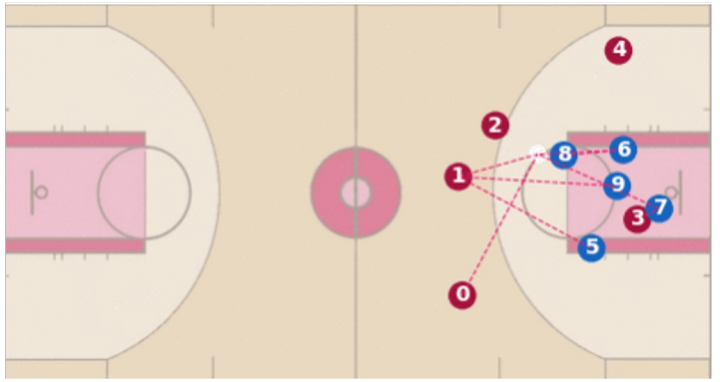
Topic samples

2. Trajectory prediction

Task definition

Trajectory Prediction is the problem of predicting the short-term (1-3 seconds) and long-term (3-5 seconds) spatial coordinates of various road-agents such as cars, buses, pedestrians, rickshaws, and animals, etc.





References

Liang, Ming, et al. "Learning lane graph representations for motion forecasting." *European Conference on Computer Vision*. Springer, Cham, 2020.

Mangalam, Karttikeya, et al. "From goals, waypoints & paths to long term human trajectory forecasting." *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2021.

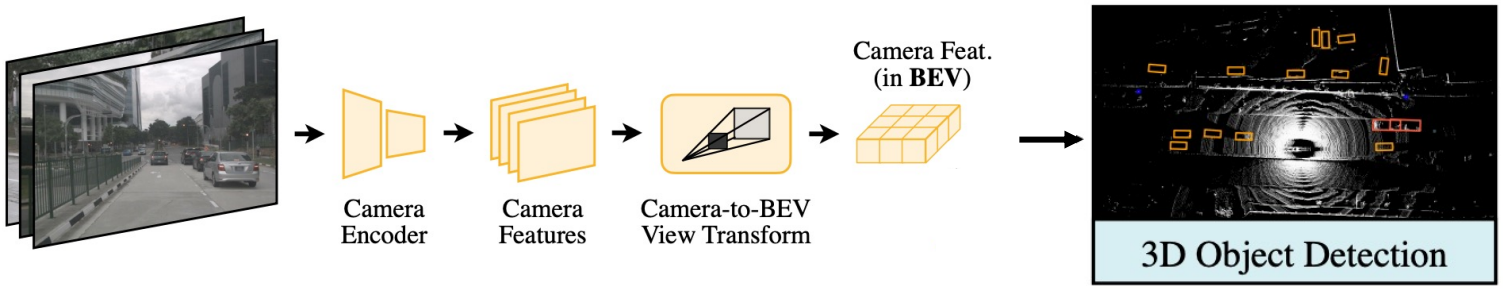
Xu, Chenxin, et al. "GroupNet: Multiscale Hypergraph Neural Networks for Trajectory Prediction with Relational Reasoning." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022.

Topic samples

3. Camera-only 3D object detection

Task definition

Monocular 3D Object Detection is the task to draw 3D bounding box around objects in a single 2D RGB image. It is localization task but without any extra information like depth or other sensors or multiple-images.



References

Reading, Cody, et al. "Categorical depth distribution network for monocular 3d object detection." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021.

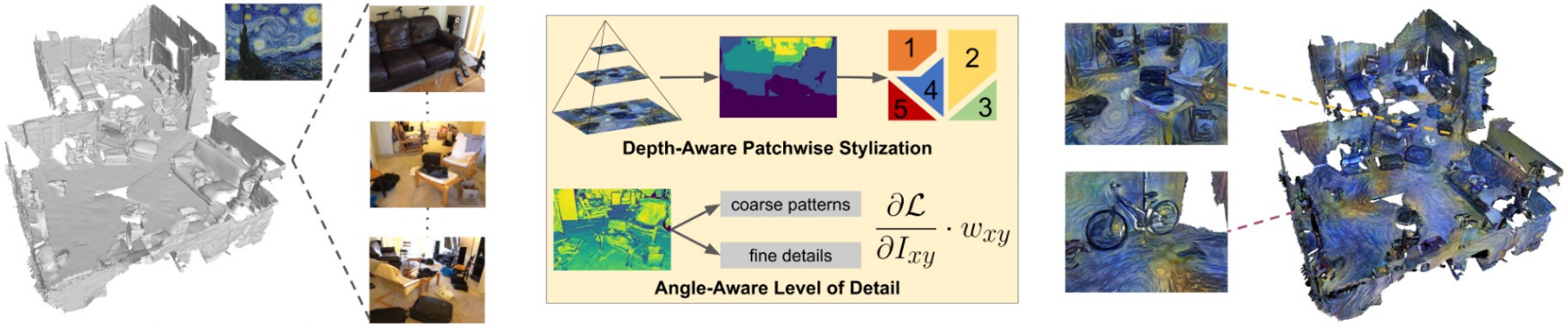
Lian, Qing, et al. "Exploring Geometric Consistency for Monocular 3D Object Detection." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022.

Topic samples

4. Style transfer

Task definition

Style transfer is the task of changing the style of an image in one domain to the style of an image in another domain.



StyleMesh optimizes a stylized texture for an indoor scene reconstruction.

References

Höllein, Lukas, Justin Johnson, and Matthias Nießner. "StyleMesh: Style Transfer for Indoor 3D Scene Reconstructions." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022.

Topic samples

5. Domain adaptation

Task definition

Domain adaptation is the task of adapting models across domains. This is motivated by the challenge where the test and training datasets fall from different data distributions due to some factor. Domain adaptation aims to build machine learning models that can be generalized into a target domain and dealing with the discrepancy across domain distributions.

References

Sun, Tao, et al. "SHIFT: A Synthetic Driving Dataset for Continuous Multi-Task Domain Adaptation." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022.

03 Writing Criteria

Related works

- Complete and clear, no grammar error.
- Present with a well organization.
- Clarify the relation between previous work and yours.

Sample

Motion prediction aims to predict future trajectories of agents conditioned on past movements.

→ Introduce this key component

Recently, three mechanisms are developed to in the motion prediction.
The first is spatial-centric mechanism, it represents the actors' trajectories in a unifying spatial domain and uses the spatial relationship to implicitly model the interaction between actors. For example, Social Conv [7] and MATF [40] leverage the spatial structure of the actors to learn the interactions; ChauffeurNet [3] and Motion Pre- diction [8] encode the trajectories of traffic actors and the scene context into bird's eye view images.
The second is social mechanism, it…
The third is graph-based mechanism, it…

→ Introduce previous works

Our work also follows the graph-based mechanism, but we add extra design from two aspects: (i) capture directed interactions; and (ii) provide interpretability for interactions.

→ Introduce your work and connections

Method

- Clearly describe the method/trick you use in detail.
- If your method is based on an existing baseline, clarify the extra design you use.

Sample

2.3. Goal selection

After the dense probability estimation, we use non-maximum suppression (NMS) algorithm to select goals. NMS iteratively selects the goal with the highest probability and removes the goals which are close to the selected goal. The first K selected goals are the predicted goals.

2.4. Trajectory completion

Similar to TNT, the last step is to complete each trajectory conditioned on the selected goals. We only have one ground truth trajectory, so we apply a teacher forcing technique [12] by feeding the ground truth goal during training. The loss term is the offset between the predicted trajectory $\hat{s}$ and the ground truth trajectory s:

$$\mathcal{L}_{\text{completion}} = \sum_{t=1}^{T} \mathcal{L}_{\text{reg}}(\hat{s}_t, s_t) \quad (5)$$

where $\mathcal{L}_{\text{reg}}$ is the smooth $l1$ loss between two points.

Introduce the several methods/tricks you used

Write the essential equations

Experiment

- Demonstrate the results in a proper form. Compare with the other baselines.

Sample

Waymo Open Dataset Motion Prediction Challenge. We evaluate the effectiveness of DenseTNT on the Waymo Open Dataset Motion Prediction Challenge. As shown in Table 2, our method ranks 1^{st} on the leaderboard. The official metric is mAP, which provides a full picture of the model performance [4]. The breakdown performance on each category is shown in Table 3.

Introduce your performance

Table 2. Top 10 entries of the Waymo Open Dataset Motion Prediction Challenge. mAP is the official ranking metric.

Method	minADE	minFDE	Miss Rate	Overlap Rate	mAP
DenseTNT 1^{st} (Ours)	1.0387	1.5514	0.1573	0.1779	**0.3281**
TVN 2^{nd}	0.7558	1.5859	0.2032	**0.1467**	0.3168
Star Platinum 3^{rd}	0.8102	1.7605	0.2341	0.1774	0.2806
SceneTransformer	**0.6117**	**1.2116**	**0.1564**	0.1473	0.2788
ReCoAt	0.7703	1.6668	0.2437	0.1642	0.2711
AIR	0.8682	1.6691	0.2333	0.1583	0.2596
SimpleCNNOnRaster	0.7400	1.4936	0.2091	0.1560	0.2136
CNN-MultiRegressor	0.8257	1.7101	0.2735	0.1640	0.1944
GOAT	0.7948	1.6838	0.2431	0.1726	0.1930
Waymo LSTM baseline	1.0065	2.3553	0.3750	0.1898	0.1756

Compare your results to the baselines in the form of tables, figures, etc.

Experiment

- Conduct ablation test to verify the effectiveness of each part of your method.
- Analyze why the components work for your problem.

Sample

Model	Description	w_{imit}	w_{env}
$\mathcal{M}_0$	Imitation with Past Dropout	1.0	0.0
$\mathcal{M}_1$	$\mathcal{M}_0$ + Traj Perturbation	1.0	0.0
$\mathcal{M}_2$	$\mathcal{M}_1$ + Environment Losses	1.0	1.0
$\mathcal{M}_3$	$\mathcal{M}_2$ with less imitation	0.5	1.0
$\mathcal{M}_4$	$\mathcal{M}_2$ with Imitation Dropout	*Dropout probability = 0.5 (see Section 5.3).*	

Table 3: Model configuration for the model ablation tests.

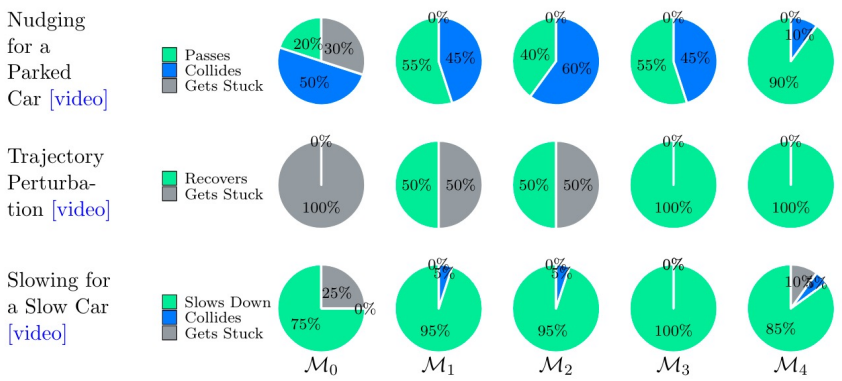


Figure 7: Model ablation test results on three scenario types.

Sample

6.3.1 Model Ablation Tests

Here, we present results from experiments using the various models in the closed-loop simulation setup. We first evaluated all the models on simple situations such as stopping for stop-signs and red traffic lights, and lane following along straight and curved roads by creating 20 scenarios for each situation, and found that *all the models worked well in these simple cases.* Therefore, we will focus below on specific complex situations that highlight the differences between these models.

Choose a proper form to demonstrate the ablation tests.

Template

- We will provide two templates for the competition track and research track. See the templates for more writing instructions. We will upload the templates to canvas.

Title of your project (Competition track)

Member 1, Member 2, Member 3, and Member 4 *

Requirement. Your final report should be written in the same style as a CV research paper. Your final report must be a PDF file. Your final report can be **maximum 8 pages** (excluding references). Consider using the following section structure, though you can use a different structure. **Delete** this paragraph in your final report.

Abstract

An abstract should concisely (less than 300 words) motivate the problem, describe your aims, describe main methods/tricks, and highlight your competition rank/scores.

1. Introduction

The introduction explains the task, why it's difficult, interesting, or important, and explains the main framework/baseline you used, explains the extra design or tricks you used. You can put one or some representative data samples like images for better understanding the task. Though an introduction covers similar material as an abstract, the introduction gives more space for motivation, detail, references to existing work, and to capture the reader's interest.

2. Related Work

This section helps the reader understand the research context of your work, by providing an overview of existing work in the area. Remember to describe the connection between your method and previous baselines. This part is similar to the requirement of review summary in the proposal report.

3. Approach

This section details your approach(es) and tricks to the competition task. For example, this is where you describe the architecture of your neural network(s), and any other key methods or useful tricks you used.

*This submission is the project proposal for course ECE4880J in 2022SU. This course is held by UM-SJTU Joint Institute, Shanghai Jiao Tong University. The mentor of this course is Dr. Siheng Chen.

4. Experiments

This section contains the following.

4.1. Data

Describe the dataset you are using (provide website). If it's not already clear, make sure the associated task is clearly described. Being precise about the exact form of the input and output can be very useful for readers attempting to understand your work, especially if you've defined your own task.

4.2. Evaluation method

Describe the evaluation metric(s) the competition use, plus any other details necessary to understand your evaluation.

4.3. Experimental details

Report how you ran your experiments **in detail**. (e.g. model configurations, learning rate, training time, hyperparameters, etc.)

4.4. Results

Report the quantitative results that you have found so far. Use a table or plot to compare your results with baseline results from other teams. If your method is based on an existing baseline, you must report its results. Comment on your quantitative results. Are they what you expected? Why better than some baselines?

4.5. Ablation studies

Report the ablation study result for every of your proposed methods/tricks. You should compare the results withwithout your proposed methods/tricks. You also should present the influence of taking different values of important hyperparameters in your method. You can use tables or plots for presenting results. You would better understand your methods/tricks and give some reasons about how they work and why they succeed.

5. Conclusion

Summarize the main findings of your project, and what you have learnt. Highlight your achievements, and note

Title of your project (Research track)

Member 1, Member 2, Member 3, and Member 4 *

Requirement. Your final report should be written in the same style as a CV research paper. Your final report must be a PDF file. Your final report can be **maximum 8 pages** (excluding references). Consider using the following section structure, though you can use a different structure. **Delete** this paragraph in your final report.

Abstract

An abstract should concisely (less than 300 words) motivate the problem, describe your aims, describe your contribution, and highlight your main idea(s).

1. Introduction

The introduction explains the problem, why it's difficult, interesting, or important, how and why current methods succeed/fail at the problem, and explains the key ideas of your approach and results. Though an introduction covers similar material as an abstract, the introduction gives more space for motivation, detail, references to existing work, and to capture the reader's interest.

2. Related Work

This section helps the reader understand the research context of your work, by providing an overview of existing work in the area. This part is similar to the requirement of review summary in the proposal report.

3. Approach

This section details your approach(es) to the problem. For example, this is where you describe the architecture of your neural network(s), and any other key methods or algorithms.

4. Experiments

This section contains the following.

*This submission is the project proposal for course ECE4880J in 2022SU. This course is held by UM-SJTU Joint Institute, Shanghai Jiao Tong University. The mentor of this course is Dr. Siheng Chen.

4.1. Data

Describe the dataset(s) you are using (provide references). If it's not already clear, make sure the associated task is clearly described. Being precise about the exact form of the input and output can be very useful for readers attempting to understand your work, especially if you've defined your own task.

4.2. Evaluation method

Describe the evaluation metric(s) you use, plus any other details necessary to understand your evaluation. Some projects will have clear metrics from prior work on given datasets, but we realize that other projects will define their own metrics. If you're defining your own metrics, be clear as to what you're hoping to measure with each evaluation method (whether quantitative or qualitative, automatic or human-defined!), and how it's defined.

4.3. Experimental details

Report how you ran your experiments (e.g. model configurations, learning rate, training time, etc.)

4.4. Results

Report the *qualitative results*. That is, compare your proposed method with previous works, and conduct ablations for every of your proposed methods/tricks. You should compare the results withwithout your proposed methods/tricks. You also should present the influence of taking different values of important hyperparameters in your method. You can use tables or plots for presenting results. You would better understand your methods/tricks and give some reasons about how they work and why they succeed.

Report the *quantitative results* that you have found so far. Use a table or plot to compare results and compare against baselines. Comment on your quantitative results. Are they what you expected? Better than you expected? Worse than you expected? Why do you think that is? What does that tell you about your approach?

5. Conclusion

Summarize the main findings of your project, and what you have learnt. Highlight your achievements, and note

04 Presentation

Detailed grading policy

Presentation (10%)

- Introduce the background (1%)
- Explain your method clearly (3%)
- Show the performance with nice DEMO (4%)
- Overall quality of the video (1%)
- Time limitation: 3~5 min (1%)

You are encouraged to present your result by every means in addition to the slides.

05

Q & A