Data Selection Matters: Towards Robust Instruction Tuning of Large Multimodal Models

Xu Yang[1] Chen Liu[1*] Ying Wei[2*] [1]City University of Hong Kong [2]Zhejiang University

Background & Motivation

- Visual instruction selection aligns LMMs with human intent efficiently.
- However, existing methods **inherit dataset biases** (e.g., position bias, spurious correlations), causing **biased model behaviors.**
- These biases lead to **drops in robustness** under simple perturbations (e.g., shuffling option order, changing option symbols).
- **Text-only catastrophic forgetting** exacerbates this vulnerability.

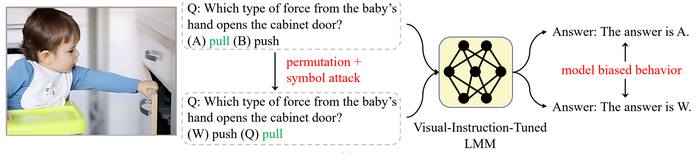

Table 1. ***Information Proxy*** indicates the representation used to compute the information measure. ***Objective*** means the selection goal emphasized when ranking samples. ***Task-Aware Selection*** denotes methods explicitly target a specific task. ***Downstream-Data-free*** marks no downstream-task samples are required.

Method	Information Proxy	Objective	Task-Aware Selection	Downstream-Data-free
LESS [107]	Gradient	Quality	✓	✗
ICONS [106]	Gradient	Quality	✓	✗
TIVE [68]	Gradient	Diversity	✓	✓
COINCIDE [51]	Feature	Diversity	✗	✓
ARDS (Ours)	Feature	Robustness	✓	✓

Contributions

- **ARDS:** A gradient-free, robustness-aware data selection framework for enhancing the robustness and data efficiency of visual instruction tuning.
- **Conversation Vector:** Represents multi-turn samples using attention-score weighted aggregation.
- **Worst-case Evaluation Subgroups:** Built via clustering and task-specific perturbations to identify high-quality training samples.
- **Robust Training Mixture**: Curated with a small model, showing strong cross-model transferability and improving robustness for a larger model.

Framework

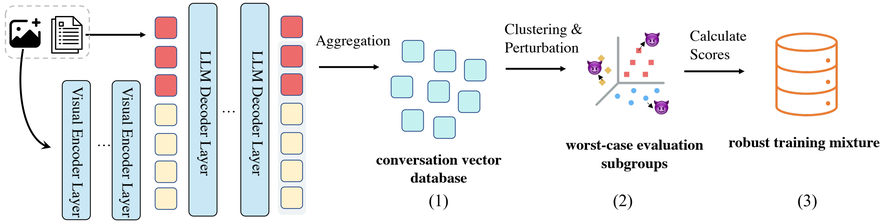

Figure 1. ARDS Pipeline. (1) Extract token embeddings for multi-turn conversations and aggregate key information using attention scores. (2) Perform hierarchical clustering in the embedding space and apply perturbations to identify samples most vulnerable to biased behaviors. (3) Measure distance to worst-case subgroups and select top-ranked samples to build the robust training mixture.

Algorithm 1 Data Selection for Robust Visual Instruction Tuning (ARDS)

Require: Training corpus $\mathcal{D}_{\text{train}}$ of size N; Proxy model f_θ; Number of subgroups K; Subgroup budget B

Step 1: Conversation Vector Database

for $i = 1$ to N **do**

  Extract token embeddings $\mathbf{H}_t$ and attention matrix $\mathbf{A}$ of last token $\mathbf{H}_L$ using f_θ

  Aggregate visual and textual token representations via attention-weighted mechanism: $\widehat{\mathbf{H}} = \sum_{t=1}^{L-1} \mathbf{A}_{L,t} \cdot \mathbf{H}_t$

  Form the conversation vector $r_i = [\mathbf{H}_L; \widehat{\mathbf{H}}]$

end for

Step 2: Worst-case Evaluation Subgroups

Cluster a held-out subset of $\mathcal{D}_{\text{train}}$ into subgroups $\{\mathcal{C}_m\}_{m=1}^{M}$ in the embedding space.

Apply perturbations and compute loss differences to identify most vulnerable samples: $\mathcal{S}_m = \text{top}_B \{\mathbf{x} \in \mathcal{C}_m : |\ell(\mathbf{x}) - \ell(\mathbf{x}')|\}$

Form the worst-case evaluation subgroups $\mathcal{S} = \{\mathcal{S}_m\}_{m=1}^{M}$

Step 3: Quality Score Measure

for $i = 1$ to N **do**

  Calculate cosine similarity: $d_{i\mathcal{S}_m} = \frac{1}{B} \sum_{j \in \mathcal{S}_m} \cos(r_i, r_{\mathcal{S}_m,j})$

  Calculate the quality score using subgroup difficulty weighting: $\mathcal{I}(x_i) = \frac{\sum_{m=1}^{M} \exp(\ell_{\mathcal{S}_m}) \cdot d_{i\mathcal{S}_m}}{\sum_{m=1}^{M} \exp(\ell_{\mathcal{S}_m})}$

end for

Select top-ranked samples by $\mathcal{I}(x_i)$ to construct $\mathcal{D}_{\text{robust}}$.

Output: Robust training mixture $\mathcal{D}_{\text{robust}}$

Experiments & Analysis

Table 1. Zero-shot robust accuracies (%, ↑) against spurious correlation and position bias.

Selection Method	Data Percentage	ScienceQA Clean	PA	SA	SA + PA	Avg.	SEED-Bench Clean	PA	SA	SA + PA	Avg.	MMBench-EN Clean	PA	SA	SA + PA	Avg.
Full	100%	69.76	54.34	65.74	37.63	56.87	59.65	41.92	54.83	22.40	44.69	74.84	61.15	69.39	41.09	61.62
Random	30%	69.76	52.60	59.44	23.75	51.39	56.84	35.74	46.58	12.73	37.97	74.20	57.75	65.49	31.83	57.32
LESS-SciQA [107]	30%	68.42	55.63	64.70	34.95	55.93	55.82	36.30	52.32	18.19	40.66	72.14	57.89	67.54	34.51	58.02
RHO-LOSS [76]	30%	64.01	36.89	59.44	21.42	45.44	53.97	25.07	48.36	11.26	34.67	70.82	49.90	66.94	32.83	55.12
COINCIDE [51]	30%	67.72	52.21	61.08	28.06	52.27	57.49	36.02	48.93	15.88	39.58	73.78	58.65	68.10	37.65	59.54
ARDS (ours)	30%	69.26	**59.40**	**68.57**	**47.60**	**61.21**	58.11	40.73	**56.83**	**31.52**	**46.80**	74.43	61.03	**72.37**	**53.22**	**65.26**

Selection Method	Data Percentage	A-OKVQA Clean	PA	SA	SA + PA	Avg.	MMMU Clean	PA	SA	SA + PA	Avg.	ARC-e Clean	PA	SA	SA + PA	Avg.
Full	100%	80.52	72.31	78.34	55.02	71.54	35.06	10.15	33.65	4.84	20.92	36.76	11.11	25.25	0.83	18.48
Random	30%	78.25	66.29	70.13	35.72	62.59	34.00	9.21	**35.77**	5.43	21.10	38.95	12.38	33.99	1.36	21.67
LESS-SciQA [107]	30%	78.60	66.72	74.41	45.94	66.42	37.43	11.81	33.53	4.49	21.82	37.86	13.57	35.18	3.03	22.41
RHO-LOSS [76]	30%	76.86	55.02	71.00	37.64	60.13	34.00	5.31	32.23	3.19	18.68	38.21	5.49	34.39	1.27	19.84
COINCIDE [51]	30%	77.55	65.59	72.66	44.10	64.97	**37.90**	9.80	33.29	3.54	21.13	38.25	11.86	36.06	2.64	22.20
ARDS (ours)	30%	78.34	71.09	77.64	**64.72**	**72.95**	37.54	**12.75**	34.24	**6.97**	**22.88**	**39.92**	**16.95**	**37.15**	**8.26**	**25.57**

▲ Achieve Largest Robustness Gains with Curated Robust Training Mixture

Table 2. Transferability across large multimodal architectures and training settings.

Proxy Model	Target Model	Selection Method	Data Percentage	ScienceQA Clean	PA	SA	SA + PA	Avg.	SEED-Bench Clean	PA	SA	SA + PA	Avg.	MMBench-EN Clean	PA	SA	SA + PA	Avg.
-	LLaVA-1.5 (13B)	**Full**	100%	71.05	57.21	64.20	37.58	57.51	61.12	43.85	56.19	23.08	46.06	76.02	64.06	71.73	47.79	64.90
-	LLaVA-1.5 (13B)	**Random**	30%	70.25	54.69	63.76	31.33	55.01	59.08	39.06	52.09	16.17	41.60	75.70	59.92	69.74	39.50	61.22
LLaVA-1.5 (7B)	LLaVA-1.5 (13B)	**ARDS (ours)**	30%	**72.58**	**60.19**	**66.14**	**41.99**	**60.22**	59.94	**43.98**	**57.58**	**30.76**	**48.07**	**76.41**	**64.24**	**72.95**	**52.60**	**66.55**

Proxy Model	Target Model	Selection Method	Data Percentage	ScienceQA Clean	PA	SA	SA + PA	Avg.	A-OKVQA Clean	PA	SA	SA + PA	Avg.	MMBench-EN Clean	PA	SA	SA + PA	Avg.
-	LLaVA-Mistral (7B)	**Full**	100%	73.03	60.78	68.32	42.79	61.23	80.00	68.38	77.99	59.21	71.39	77.04	62.05	73.30	47.05	64.86
-	LLaVA-Mistral (7B)	**Random**	30%	73.08	56.22	58.70	21.17	52.29	77.47	61.31	72.93	39.21	62.73	75.31	58.51	67.48	32.87	58.54
LLaVA-1.5 (7B)	LLaVA-Mistral (7B)	**ARDS**	30%	72.04	61.77	69.16	55.53	**64.63**	81.66	72.58	80.52	69.00	**75.94**	76.97	65.37	75.17	55.19	**68.18**
-	Qwen2.5-VL (7B)	-	-	77.05	63.71	67.08	33.71	60.38	82.18	67.34	75.90	41.48	66.72	71.31	52.48	72.14	35.16	57.77
-	Qwen2.5-VL (7B)	**Random**	30%	80.32	69.31	67.43	31.78	62.21	84.54	73.01	75.90	38.25	67.92	74.27	57.36	73.83	34.63	60.02
LLaVA-1.5 (7B)	Qwen2.5-VL (7B)	**ARDS**	30%	83.84	76.55	70.15	36.19	**66.68**	85.85	77.03	77.55	42.01	**70.61**	80.85	69.81	75.44	40.29	**66.60**

▲ Strong Cross-Model Generalization Across Visual Instruction Tuning and Post-training Settings

Table 3. Results on visual mathematical reasoning benchmarks.

Selection Method	Data Percentage	MathVista Clean	PA	SA	SA + PA	Avg.	DynaMath Clean	PA	SA	SA + PA	Avg.
Full	100%	40.37	16.48	34.44	2.22	23.38	39.53	19.80	36.88	1.74	24.48
Random	30%	39.44	15.56	23.15	1.30	19.86	**38.39**	16.97	27.32	1.56	21.06
LESS [107]	30%	36.85	19.44	31.85	5.37	23.38	35.56	**20.88**	36.28	8.06	25.19
COINCIDE [51]	30%	37.41	11.48	25.00	2.04	18.98	35.74	14.80	26.90	2.47	19.98
ARDS (ours)	30%	**39.81**	**20.93**	**32.78**	**6.48**	**25.00**	36.40	20.22	**36.76**	**11.19**	**26.14**

▲ Generalization to More Challenging OOD Tasks

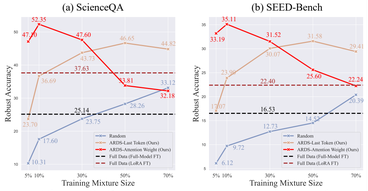

Figure 2. Robust accuracy across data scales

▼ Ablation Study

Table 4. Results comparing conversation vector variants.

Conversation Vector	Data Percentage	ScienceQA Clean	PA	SA	SA + PA	Avg.	SEED-Bench Clean	PA	SA	SA + PA	Avg.	MMBench-EN Clean	PA	SA	SA + PA	Avg.	A-OKVQA Clean	PA	SA	SA + PA	Avg.
Last Token	10%	66.68	54.04	64.06	36.69	55.36	55.13	35.02	52.29	23.90	41.58	73.11	57.61	71.06	49.53	62.82	76.33	64.45	74.59	54.32	67.42
Attention Weight	10%	69.66	55.88	69.21	52.35	61.78	53.86	38.30	53.38	35.11	45.16	72.07	59.90	71.56	57.13	65.16	77.90	70.04	76.94	66.72	72.90
Last Token	30%	69.41	57.36	65.99	43.73	59.12	58.23	41.47	56.12	30.07	46.47	75.84	62.09	73.48	53.22	66.15	78.95	71.18	78.08	63.32	72.88
Attention Weight	30%	69.26	59.40	68.57	47.60	61.21	58.11	40.73	56.83	31.52	46.80	74.43	61.03	72.37	53.22	65.26	78.34	71.09	77.64	64.72	72.95

Table 5. Results comparing different components for worst-case evaluation subgroups.

Worst-case Evaluation Perturbation	Subgroup Clustering	Data Percentage	ScienceQA Clean	PA	SA	SA + PA	Avg.	SEED-Bench Clean	PA	SA	SA + PA	Avg.	MMBench-EN Clean	PA	SA	SA + PA	Avg.	A-OKVQA Clean	PA	SA	SA + PA	Avg.
✗	✗	30%	65.29	51.66	62.17	30.44	52.39	56.75	35.40	51.01	18.56	40.43	73.76	57.80	69.14	41.44	60.53	77.12	65.33	74.06	47.60	66.03
✓	✗	30%	67.43	54.34	64.35	36.49	55.65	**58.38**	40.42	56.24	26.57	45.40	74.15	60.89	71.96	49.36	64.09	**79.04**	70.92	76.77	59.56	71.57
✓	✓	30%	**69.26**	**59.40**	**68.57**	**47.60**	**61.21**	58.11	**40.73**	**56.83**	**31.52**	**46.80**	**74.43**	**61.03**	**72.37**	**53.22**	**65.26**	78.34	**71.09**	**77.64**	**64.72**	**72.95**