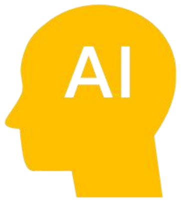
MAGR: Manifold-Aligned Graph Regularization for Continual Action Quality Assessment

Kanglei Zhou[1,3], Liyuan Wang[2], Xingxing Zhang[2], Hubert P. H. Shum[3], Frederick W. B. Li[3], Jianguo Li[4], Xiaohui Liang[1]

ECCV
EUROPEAN CONFERENCE ON COMPUTER VISION
MILANO
2024

Background – Action Quality Assessment (AQA)

- AQA aims to evaluate the quantitative performance of performed actions.



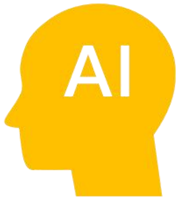
AQA system

Background – Significance of AQA

- AQA aims to evaluate the quantitative performance of performed actions.
- Mitigating human judges' biases.

Human judge

VS



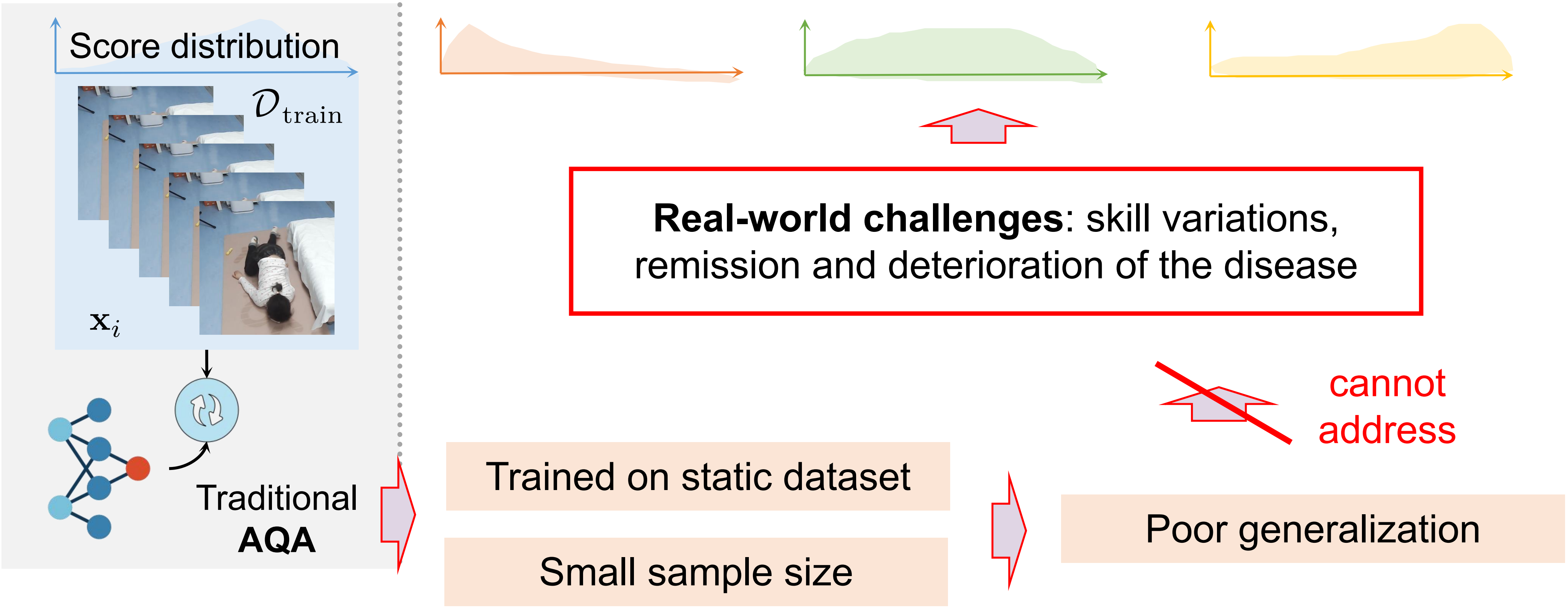
AQA system

Background – Wide AQA Application Domains

- AQA aims to evaluate the quantitative performance of performed actions.
- Mitigating human judges' biases.
- Widely used in sports, medical care, etc.

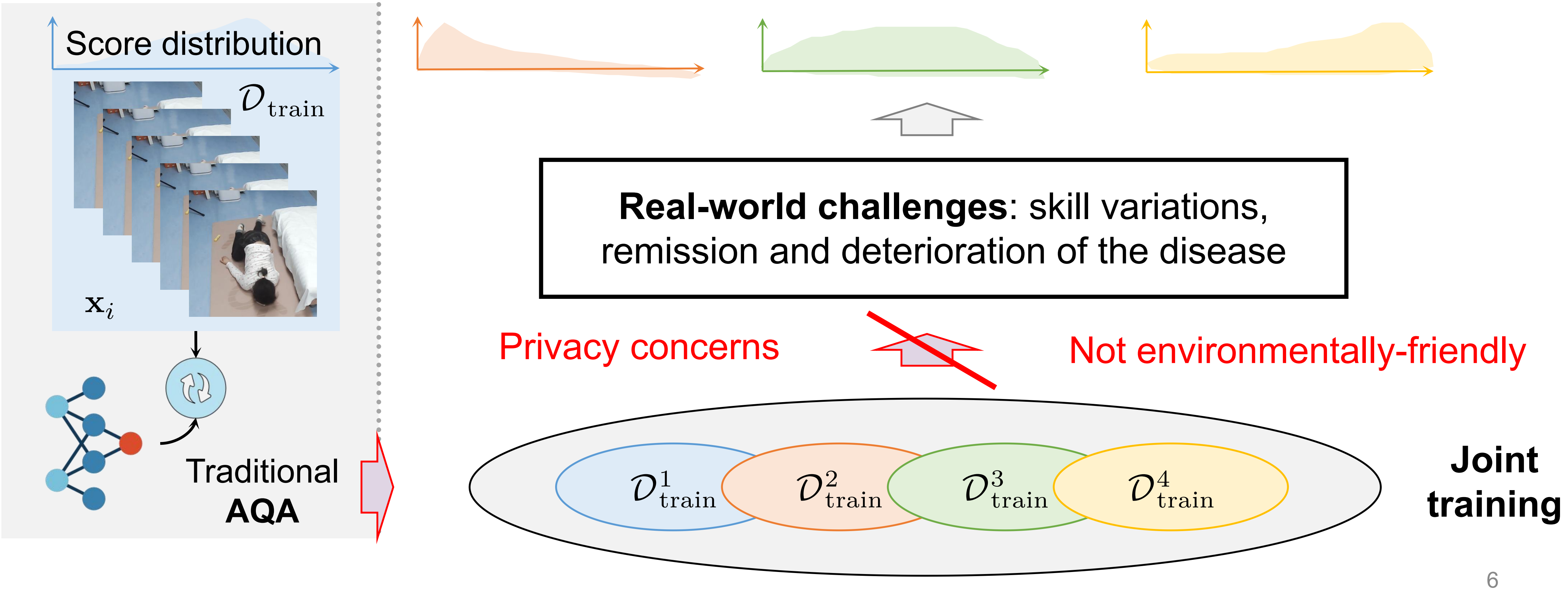
Issues with Traditional AQA Methods

- Cannot adapt to dynamically evolving changes...



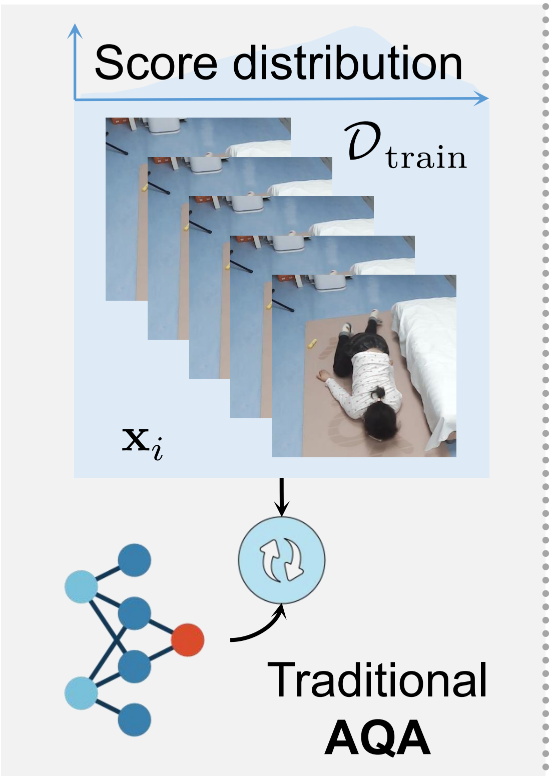
Why Traditional AQA Methods Cannot?

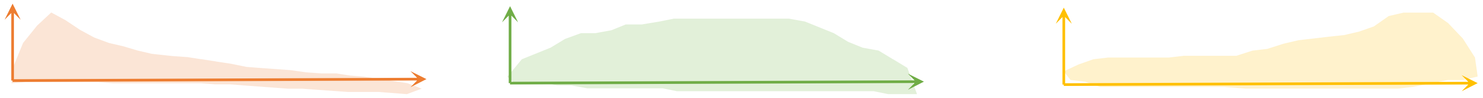
- Cannot adapt to dynamically evolving changes...



Potential Solution – Continual Learning

- Cannot adapt to dynamically evolving changes...



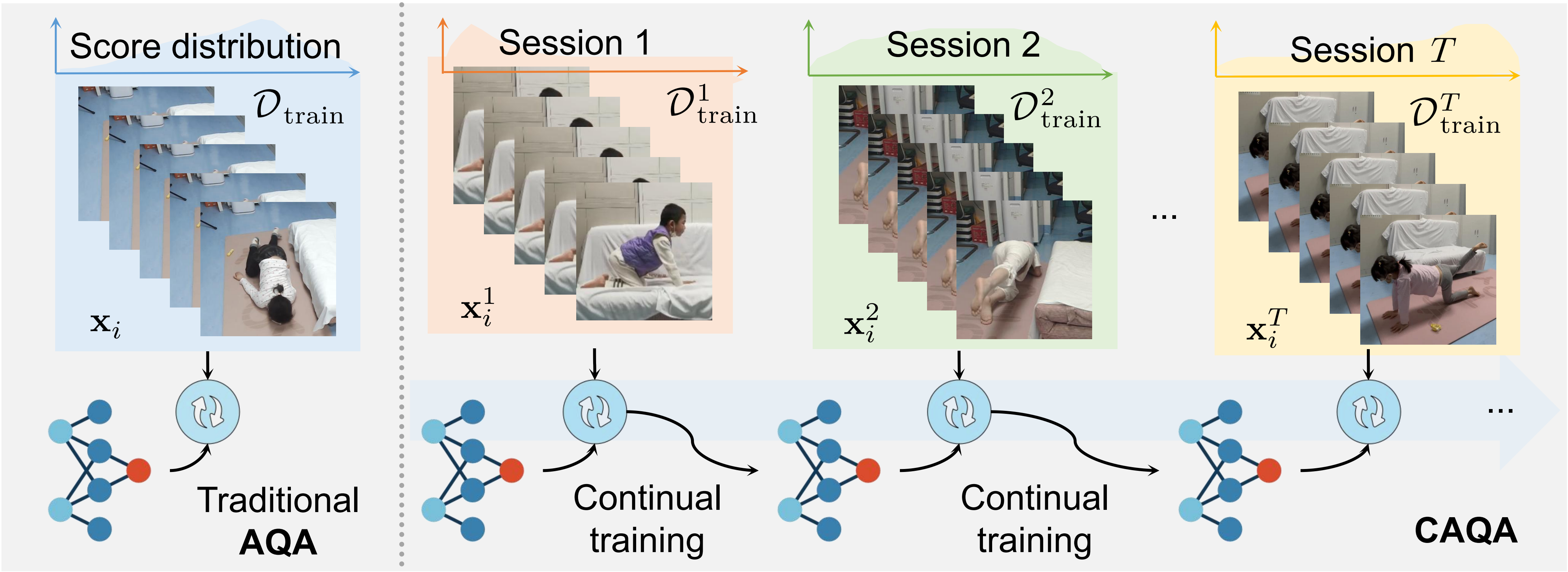


Real-world challenges: skill variations, remission and deterioration of the disease

Potential solution: Continual Learning (CL) emulates how human learns

New Task – Continual AQA (**Contribution 1**)

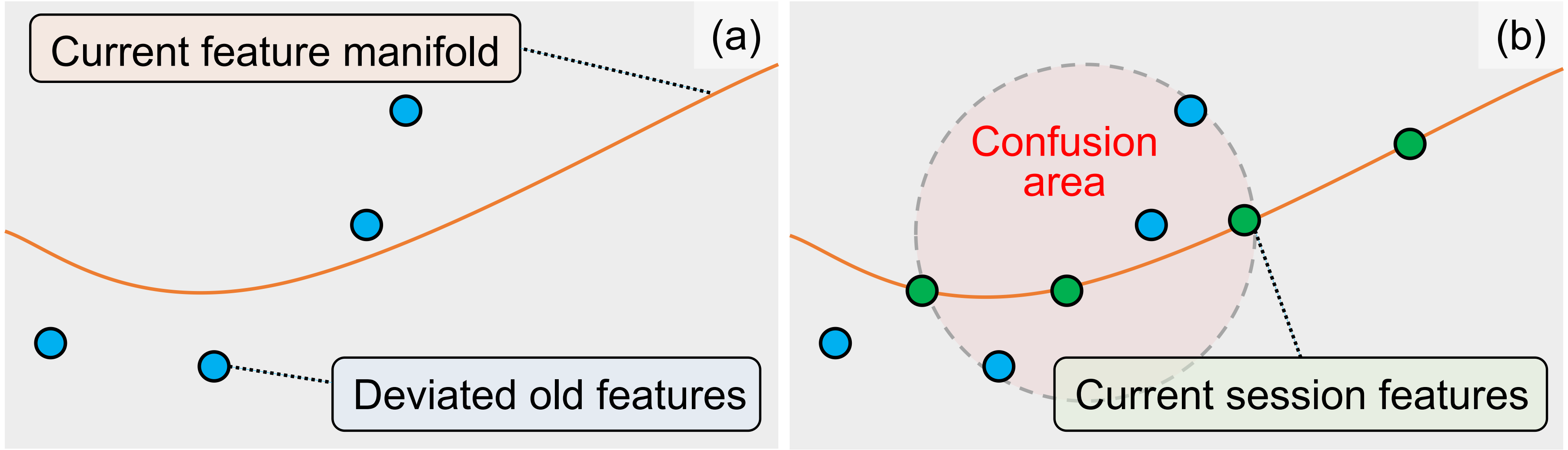
- Integrate Continual Learning (CL) into the AQA framework, protecting **user privacy** and mitigating severe forgetting.



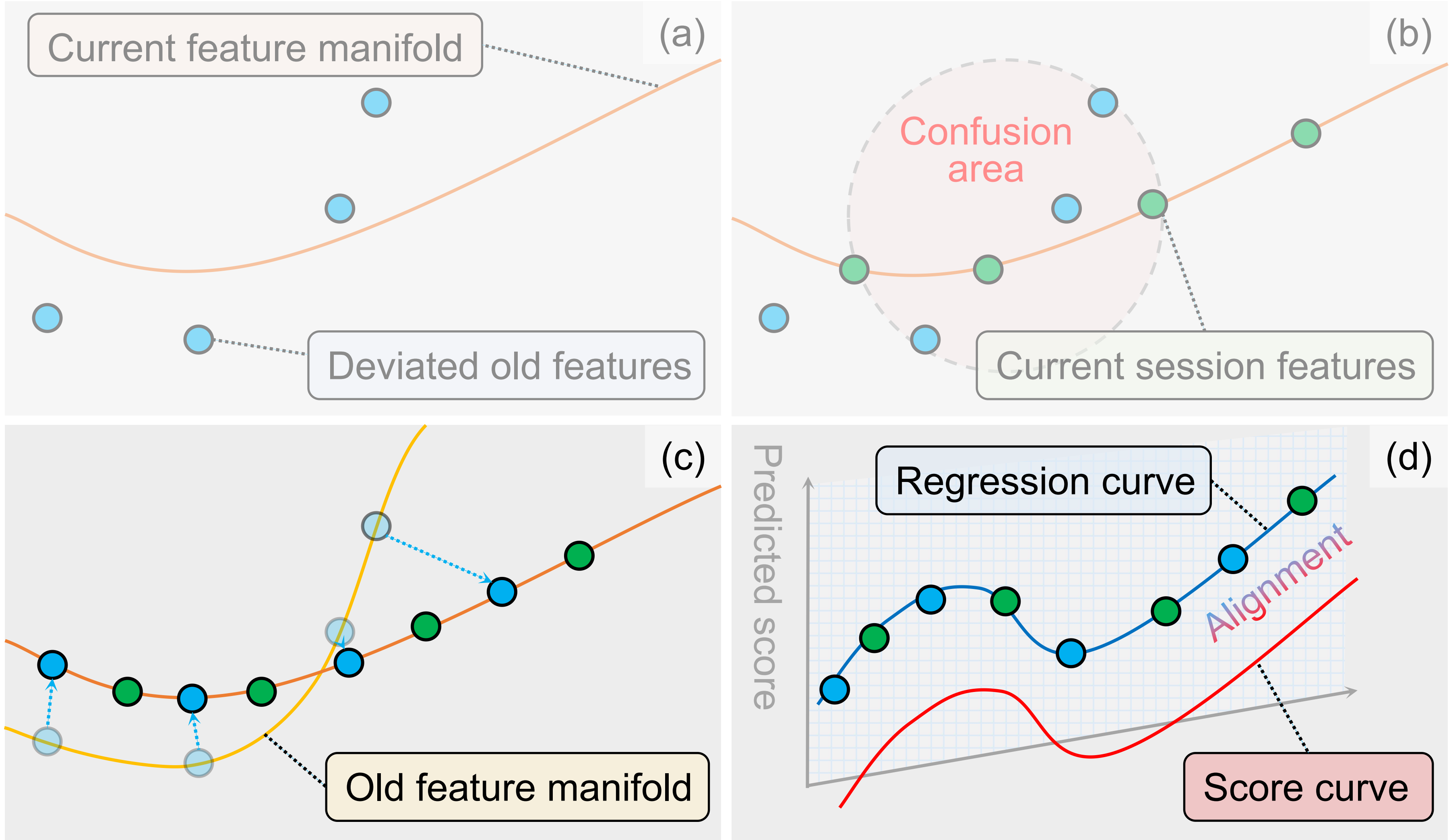
How We Takle AQA in Continual Learning

- Memory replay: an effective strategy for mitigating catastrophic forgetting
- Feature replay rather than raw data: prioritizing user privacy
- Refining the backbone to improve adaptability

Unique Challenges within CAQA – Misalignment

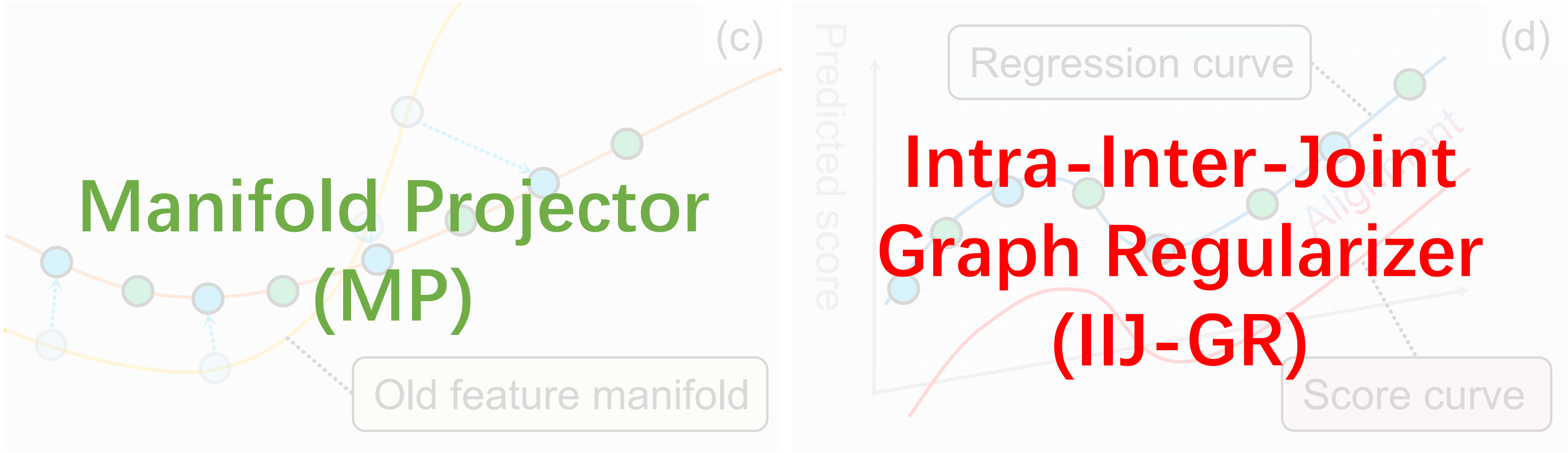


Our Solution to Misalignment (**Contribution 2**)



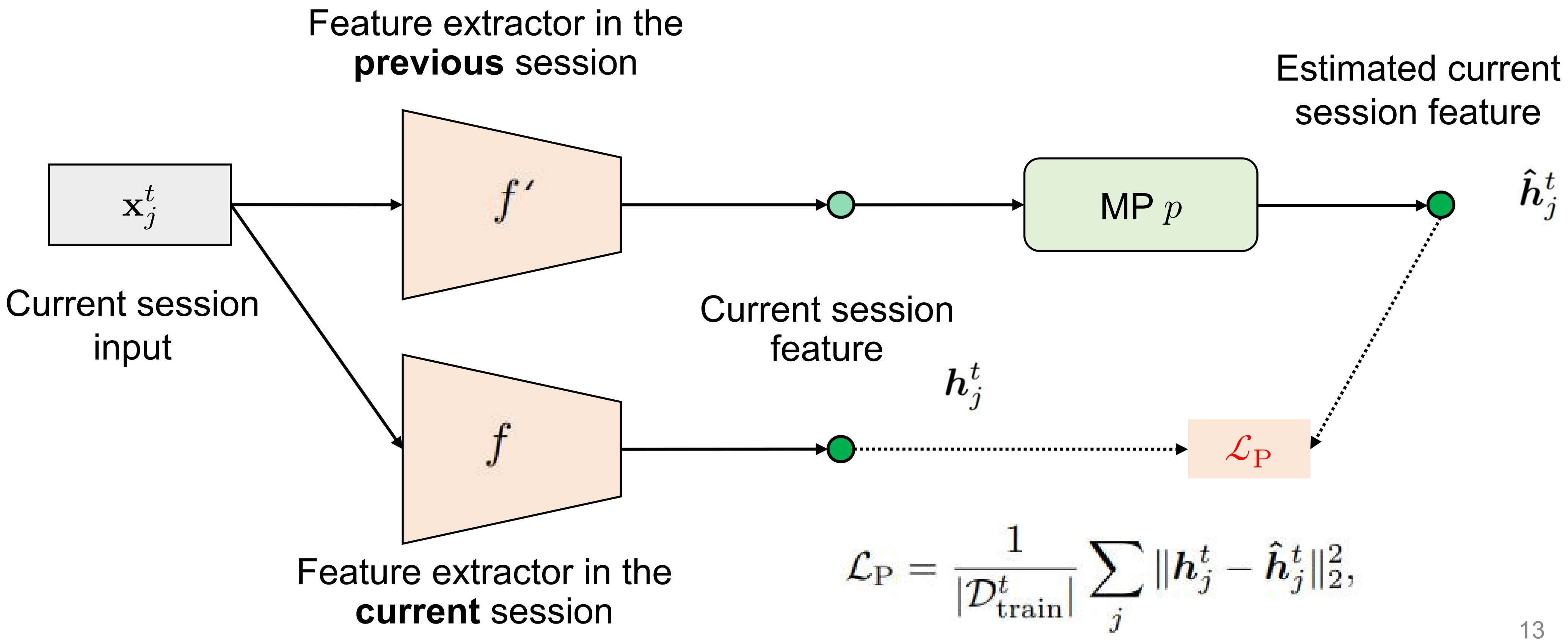
MAGR (**Implementation of Contribution 2**)

- Manifold Porjector (MP) translates old features to the current manifold
- IIJ-GR regulates the entire feature space to align with the quality space



MAGR – MP for Learning Manifold Shift

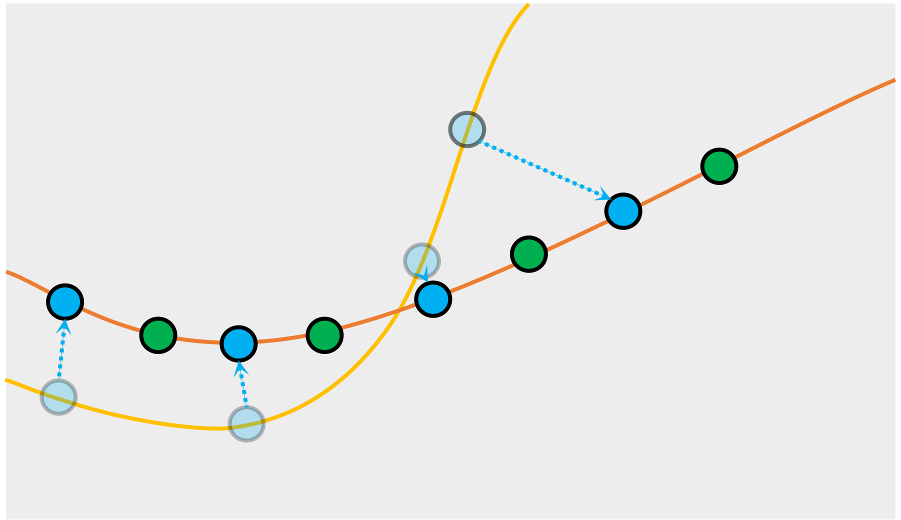
- Learn the manifold shift at each model update using dependencies from the current session data.

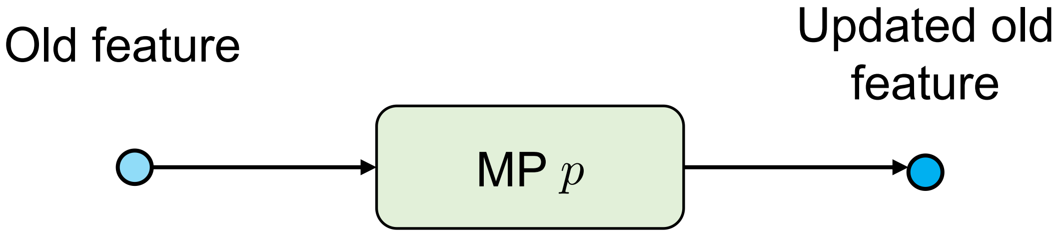


$$\mathcal{L}_{\mathrm{P}} = \frac{1}{|\mathcal{D}_{\mathrm{train}}^t|} \sum_j \|\boldsymbol{h}_j^t - \hat{\boldsymbol{h}}_j^t\|_2^2,$$

MAGR – MP for Deviated Feature Translation

- Learn the manifold shift at each model update using dependencies from the current session data.
- Translate old features to the current manifold.

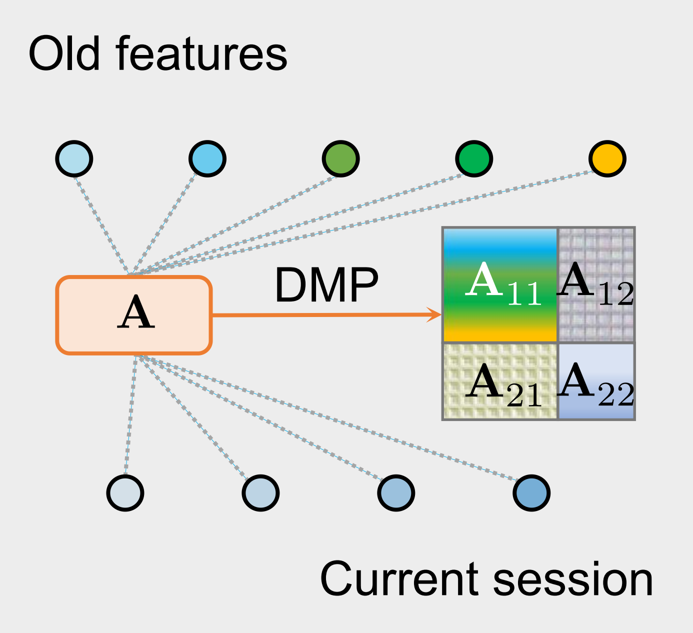




$$\tilde{\boldsymbol{h}}_i^s = \tilde{\boldsymbol{h}}_i^s + p(\tilde{\boldsymbol{h}}_i^s), \text{ where } s = 1, 2, \cdots, t-1.$$

MAGR – IIJ-GR for Feature Distribution Alignment

- IIJ-GR regulates the feature space for accurate score regression.



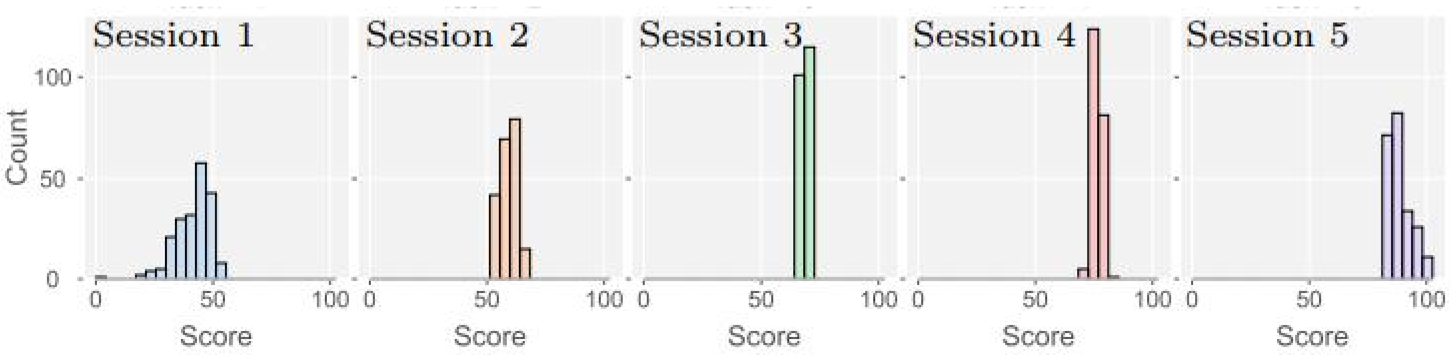
$$\mathbf{A} = \arccos(\tilde{\mathbf{H}}\tilde{\mathbf{H}}^\top), \text{ where } \tilde{\mathbf{H}} = \mathbf{H}/\|\mathbf{H}\|.$$

- $\mathbf{A}_{11}$: intra session relation
- $\mathbf{A}_{12}$: inter session relation
- $\mathbf{A}_{21}$: inter session relation
- $\mathbf{A}_{23}$: intra session relation
- $\mathbf{A}$: **joint** session relation

$$\mathcal{L}_\mathrm{R} = \boxed{L(\mathbf{A}, \mathbf{S})} + \sum_{i=1}^{2}\sum_{j=1}^{2} L(\mathbf{A}_{ij}, \mathbf{S}_{ij}),$$

Benchmark (**Contribution 3**): Data Split

- Discretizing the continuous quality space into distinct intervals corresponding to different action grades and ensuring an equal number of samples in each session, resulting in more challenging score variations.



Benchmark (**Contribution 3**): Metric

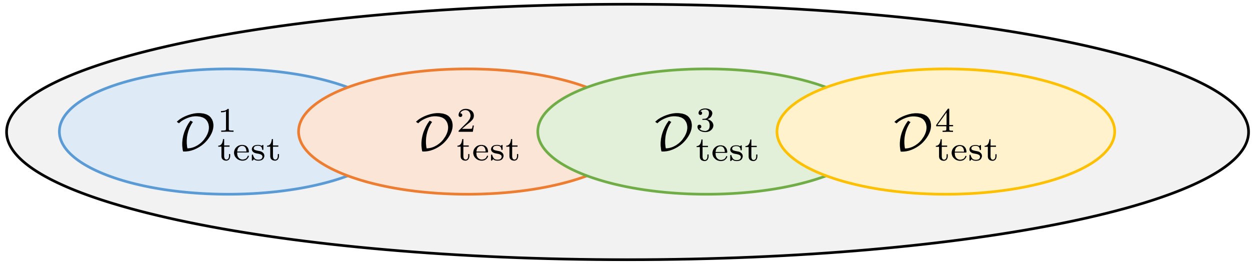
- SRCC, forgetting, forward transfer

Primary metric

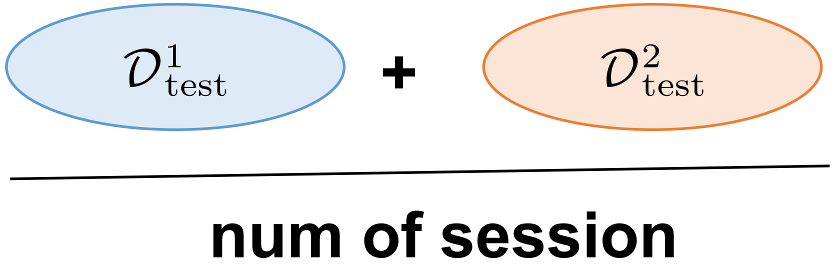
$$\rho = \frac{\sum_i (p_i - \bar{p})(q_i - \bar{q})}{\sqrt{\sum_i (p_i - \bar{p})^2 \sum_i (q_i - \bar{q})^2}},$$

Sensitive to the sample size

Testing ☑



Testing ✖



Comparison with Recent Strong Baselines

- MAGR outperforms recent strong baselines with up to 6.56%, 5.66%, 15.64%, and 9.05% correlation gains on the MTL-AQA, FineDiving, UNLV-Dive, and JDM-MSA split datasets, respectively.

Method	Publisher	Memory	MTL-AQA			FineDiving			UNLV-Dive			JDM-MSA		
			ρ_{avg} (↑)	ρ_{aft} (↓)	ρ_{fwt} (↑)	ρ_{avg} (↑)	ρ_{aft} (↓)	ρ_{fwt} (↑)	ρ_{avg} (↑)	ρ_{aft} (↓)	ρ_{fwt} (↑)	ρ_{avg} (↑)	ρ_{aft} (↓)	ρ_{fwt} (↑)
Joint Training	-	None	0.9360	-	-	0.9075	-	-	0.8460	-	-	0.7556	-	-
Sequential FT	-	None	0.5458	0.1524	0.0538	0.7420	0.1322	0.2135	0.6307	0.2135	**0.3595**	0.5080	0.1029	0.5431
SI [42]	ICML'17	None	0.5526	0.2677	0.0350	0.6863	0.2330	0.1938	0.1519	0.3822	0.0220	0.4804	0.2198	0.5431
EWC [11]	PNAS'17	None	0.2312	0.1553	0.0343	0.5311	0.3177	0.1776	0.4096	0.2576	0.3039	0.3889	0.1690	0.3120
LwF [13]	TPAMI'17	None	0.4581	0.1894	0.0490	0.7648	0.0807	0.2894	0.6081	0.1578	0.3230	0.6441	0.1127	0.2423
MER [22]	ICLR'19	Raw Data	0.8720	0.1303	0.0625	0.8276	0.1446	0.2806	0.7397	0.1321	0.0465	0.6689	**0.0635**	0.3841
DER++ [3]	NeurIPS'20	Raw Data	0.8334	0.1775	0.0433	0.8285	0.1523	0.2851	0.7206	0.1382	-0.1773	0.5364	0.0835	0.5759
TOPIC [26]	CVPR'20	Raw Data	0.7693	0.1427	0.1391	0.8006	0.1344	0.2744	0.4085	0.2647	0.1132	0.6575	0.2184	0.5492
GEM [12]	ICCV'21	Raw Data	0.8583	0.0950	0.1429	0.8309	0.0721	0.2883	0.6538	0.2322	0.0270	0.6084	0.0499	0.3566
Feature MER	-	Feature	0.7283	0.2255	0.0535	0.4914	0.2354	0.2344	0.5675	0.1322	0.1558	0.6295	0.1597	**0.6446**
SLCA [43]	ICCV'23	Feature	0.7223	0.1852	0.1665	0.8130	0.0920	0.2453	0.5551	0.1085	0.3200	0.6173	0.1705	0.4457
NC-FSCIL [39]	ICLR'23	Feature	0.8426	0.1146	0.0718	0.8087	0.0203	**0.3404**	0.6458	**0.0637**	-0.1677	0.6571	0.1295	0.4957
MAGR (Ours)	-	Feature	**0.8979**	**0.0223**	**0.1914**	**0.8580**	**0.0167**	0.2952	**0.7668**	0.0827	0.1227	**0.7166**	0.1069	0.4957

Visualizations for Addressing the Misalignment

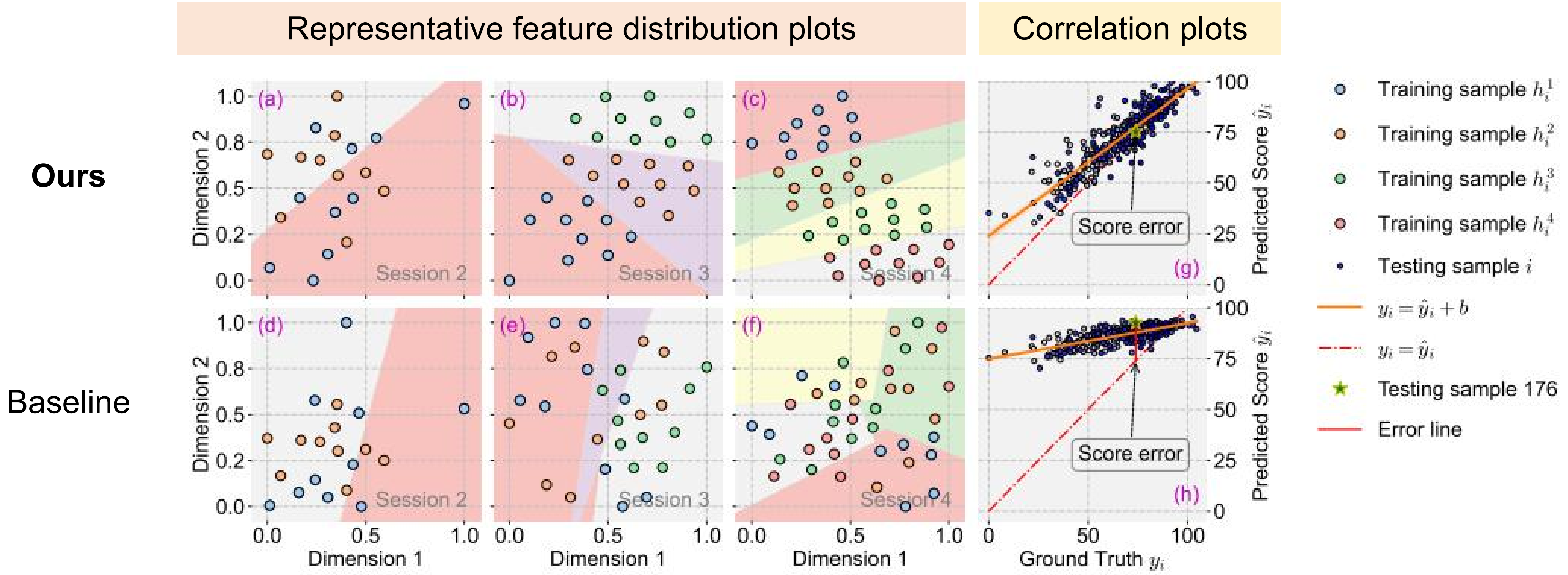


Fig. 9: Visualizations of feature distribution (a-f) and score correlation (g-h): MAGR (top) and Feature MER (bottom). The explicit division of different sessions validates the effectiveness of MAGR in mitigating catastrophic forgetting.

Ablation Study for the Effectiveness of MP

- On the MTL-AQA split dataset

MP

IIJ-GR

Setting	ρ_{avg} (↑)	ρ_{aft} (↓)	ρ_{fwt} (↑)
MAGR (Ours)	0.8979	0.0223	0.1914
w/o MP	0.6949 $^{\downarrow 23\%}$	0.1325 $^{\uparrow 494\%}$	0.0814 $^{\downarrow 57\%}$
w/o MP's residual link	0.8391 $^{\downarrow 7\%}$	0.0232 $^{\uparrow 4\%}$	0.1743 $^{\downarrow 9\%}$
w/o II-GR	0.8463 $^{\downarrow 6\%}$	0.0970 $^{\uparrow 335\%}$	0.1062 $^{\downarrow 45\%}$
w/o J-GR	0.7839 $^{\downarrow 13\%}$	0.1053 $^{\uparrow 372\%}$	0.1005 $^{\downarrow 48\%}$
w/o IIJ-GR	0.7362 $^{\downarrow 18\%}$	0.1232 $^{\uparrow 452\%}$	0.0883 $^{\downarrow 54\%}$
w/o KL (MSE loss)	0.8447 $^{\downarrow 6\%}$	0.0265 $^{\uparrow 16\%}$	0.1890 $^{\downarrow 1\%}$
w/o OUS (random sampling)	0.8619 $^{\downarrow 4\%}$	0.0876 $^{\uparrow 293\%}$	0.1027 $^{\downarrow 46\%}$

Ablation Study for the Effectiveness of IIJ-GR

- On the MTL-AQA split dataset

MP

IIJ-GR

Setting	ρ_{avg} (↑)	ρ_{aft} (↓)	ρ_{fwt} (↑)
MAGR (Ours)	0.8979	0.0223	0.1914
w/o MP	0.6949 ↓23%	0.1325 ↑494%	0.0814 ↓57%
w/o MP's residual link	0.8391 ↓7%	0.0232 ↑4%	0.1743 ↓9%
w/o II-GR	0.8463 ↓6%	0.0970 ↑335%	0.1062 ↓45%
w/o J-GR	0.7839 ↓13%	0.1053 ↑372%	0.1005 ↓48%
w/o IIJ-GR	0.7362 ↓18%	0.1232 ↑452%	0.0883 ↓54%
w/o KL (MSE loss)	0.8447 ↓6%	0.0265 ↑16%	0.1890 ↓1%
w/o OUS (random sampling)	0.8619 ↓4%	0.0876 ↑293%	0.1027 ↓46%

More Results – Robustness Analysis

Table 3: Statistics of feature deviations and correlation gains.

Dataset	FineDiving	MTL-AQA	UNLV-Dive
Degree of Feature Deviations (MSE)	26.85	35.28	51.75
Overall Correlation Gains (%)	5.66	6.56	15.64

Conclusions and Future Work

- Enable AQA to adapt to **non-stationary varitions** in real-world senarios
- Contributions**:**
  1. Define the new **CAQA** task
  2. Propose **MAGR** as **a novel solution**
  3. Establish a comprehensive **benchmark**
- Future work:
  1. Advanced network architectures like ViT
  2. Incorporate prompt-based techniques

Thanks!

Poster Location: **#148**
Time: Thu **3 Oct** 4:30 p.m. – 6:30 p.m. CEST