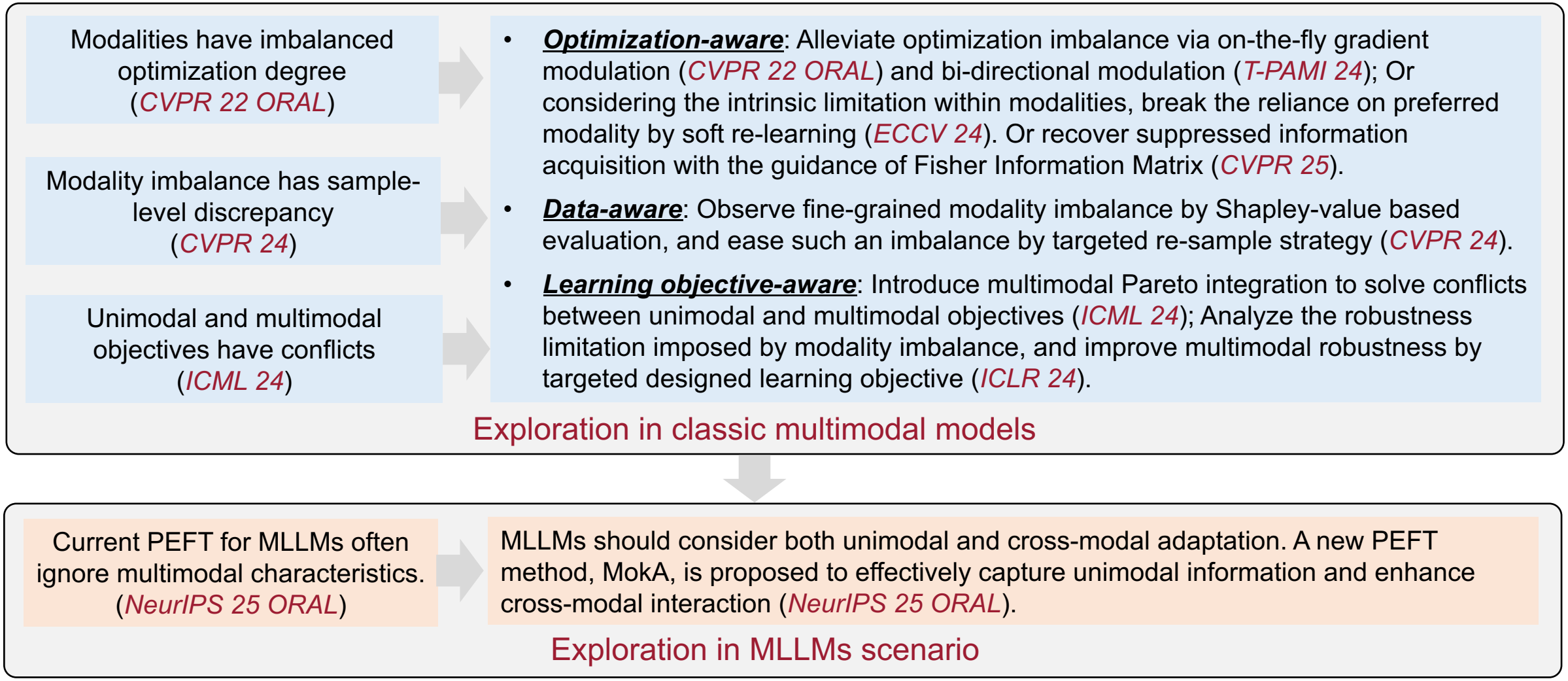
Modalities have imbalanced optimization degree (CVPR 22 ORAL)
Modality imbalance has sample-level discrepancy (CVPR 24)
Unimodal and multimodal objectives have conflicts (ICML 24)
• Optimization-aware: Alleviate optimization imbalance via on-the-fly gradient modulation (CVPR 22 ORAL) and bi-directional modulation (T-PAMI 24); Or considering the intrinsic limitation within modalities, break the reliance on preferred modality by soft re-learning (ECCV 24). Or recover suppressed information acquisition with the guidance of Fisher Information Matrix (CVPR 25).
• Data-aware: Observe fine-grained modality imbalance by Shapley-value based evaluation, and ease such an imbalance by targeted re-sample strategy (CVPR 24).
• Learning objective-aware: Introduce multimodal Pareto integration to solve conflicts between unimodal and multimodal objectives (ICML 24); Analyze the robustness limitation imposed by modality imbalance, and improve multimodal robustness by targeted designed learning objective (ICLR 24).
Exploration in classic multimodal models
Current PEFT for MLLMs often ignore multimodal characteristics. (NeurIPS 25 ORAL)
MLLMs should consider both unimodal and cross-modal adaptation. A new PEFT method, MokA, is proposed to effectively capture unimodal information and enhance cross-modal interaction (NeurIPS 25 ORAL).
Exploration in MLLMs scenario