PERSPECTIVE-CONDITIONED MULTIMODAL REASONING FOR E-COMMERCE DISPUTE VERDICTS

Anonymous authors
Paper under double-blind review

ABSTRACT

The growing volume of E-commerce disputes creates an increasing demand for automated verdict systems. Existing multi-agent methods introduce diverse role prompts and interaction contexts, aiming to elicit complementary reasoning and improve verdict accuracy. However, their effectiveness remains limited by the inherent biases of underlying models. To address this, we propose *PrismJudge*, a Perspective-Conditioned Multimodal Reasoning Framework for E-Commerce Dispute Verdict (EDV). Inspired by diverse human perspectives, *PrismJudge* constructs K perspective branches, each associated with a learnable latent perspective condition. These conditions jointly guide evidence grounding and verdict reasoning to capture diverse responsibility attribution patterns. Specifically, *Perspective-Conditioned Visual Cue Grounding* (PCVCG) exploits perspective-conditioned visual scores to selectively ground informative visual tokens, enabling the discovery of pivotal cues from redundant inputs. Built upon this, *Perspective-Conditioned Verdict Rationale Generation* (PCVRG) injects perspective conditions into the verdict reasoning process through case-adaptive perspective states and perspective-conditioned LoRA, enabling branches to share common verdict knowledge while learning distinct reasoning viewpoints for the final verdicts. Subsequently, *Rationale-Aligned Joint Learning* (RAJL) aligns human rationales with perspective branches and jointly optimizes generation, rationale matching, and diversity objectives to encourage diverse cue interpretation and complementary verdict generation. Extensive experiments show that *PrismJudge* improves accuracy by $\Delta_{\mathrm{Acc}}8.58\%$ and reduces the buyer–seller accuracy gap by $\Delta_{\mathrm{Gap}}7.24\%$ compared with the strongest baseline, highlighting the effectiveness of perspective-conditioned reasoning for accurate and balanced EDV.

1 INTRODUCTION

With the rapid growth of E-commerce transactions, the volume of online disputes has increased substantially, creating an urgent demand for efficient dispute resolution (Tsurel et al., 2020; Chen et al., 2023). To maintain platform credibility, E-commerce platforms have introduced a crowd-sourced jury mechanism for dispute verdicts: Disputants provide multimodal evidence to support their claims through multiple rounds of submissions. Meanwhile, a panel of human jurors identify dispute-related cues from the submitted evidence, and provide verdicts with corresponding rationales. However, the human-intensive nature of the mechanism incurs considerable cost and latency, hindering its scalability in large-scale dispute verdicts (Sun et al., 2026). Therefore, developing intelligent systems for dispute verdicts has become an urgent practical need.

With the emergence of MLLMs, multi-agent systems (MAS) provide a promising direction for building efficient, reliable, and automated verdict methods (Chan et al., 2024; He et al., 2024; Du et al., 2024). They typically assign different roles to LLM agents, such as judges and lawyers, prompting each agent to clarify the relationship between evidence and dispute focus, thereby deriving an initial decision. Subsequently, agents interact with other role-specific agents to refine their decisions and reach the final verdict (Chen et al., 2025; Jia et al., 2025). By introducing diverse roles and interaction contexts, these MAS methods aim to improve verdict accuracy through complementary perspectives, as shown in Fig. 1. However, *role diversity alone does not guarantee complementary decision perspectives.* On the one hand, existing studies have shown that LLM-based agents remain influenced by the inherent biases of underlying foundation models, and role-based prompts are insufficient to elicit genuinely diverse viewpoints of the simulated agents

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

(Wang et al., 2025; Peng et al., 2026). Rather, they still tend to produce responses aligned with collective stereotypes (Taubenfeld et al., 2024). On the other hand, our empirical evaluation in Fig. 3(c) reveals that existing LLMs and MLLMs exhibit substantial prediction bias, with a systematic tendency toward buyer-favoring verdicts. For example, *Gemini-2.5-Flash-Lite* exhibits a pronounced performance asymmetry across verdict categories, achieving 81.66% accuracy on buyer-win cases but only 35.85% on seller-win, resulting in a 45.81% disparity. Meanwhile, existing role-prompting approaches remain insufficient in alleviating verdict bias, as they are still constrained by the single reasoning pattern of foundation models. For instance, *CyberJurors* still exhibits a 22.98% accuracy gap between these two categories. **Therefore, mitigating the inherent verdict bias is crucial for intelligent verdict systems.**

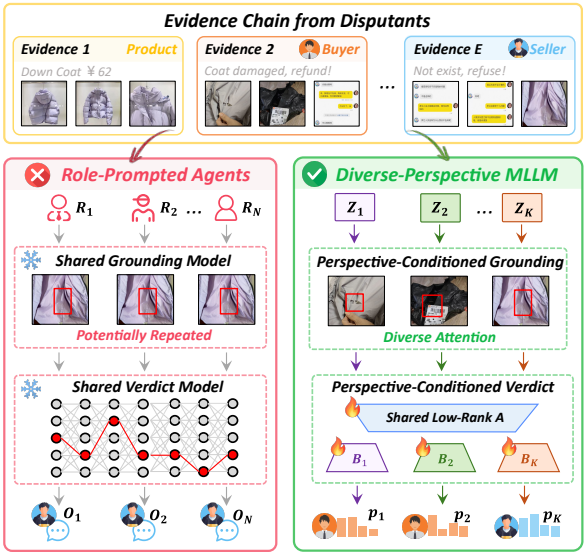


Figure 1: Comparison between role prompting and perspective-conditioned reasoning. *PrismJudge* learns diverse perspective conditions for complementary cue grounding and verdict generation.

"A mountain appears different from different viewpoints."

Inspired by the above observations, a natural solution is to introduce a perspective-guided verdict model that learns diverse reasoning patterns, thereby reducing reliance on a single perspective. However, this endeavor presents two significant challenges. First, disputants often submit large amounts of multimodal evidence to support their respective claims, while the pivotal cues that determine the verdict decisions are frequently subtle, such as product stains or minor defects. Meanwhile, human jurors do not review all evidence indiscriminately. Instead, they selectively attend to informative visual cues based on their visual salience (Mischkowski et al., 2021; De Haas et al., 2019). *Therefore, it is necessary to develop a perspective-aware mechanism that grounds dispute-relevant visual cues from redundant evidence.* Second, verdict-related perspectives influence how humans interpret visual cues in relation to the focus of a dispute, consequently affecting responsibility attribution (Kahan et al., 2008; Pennington & Hastie, 1986). For example, from different perspectives, product stains may indicate either product description inconsistency or normal wear of a second-hand item. However, merely modifying the visual cue extraction process may be insufficient to alter the model's underlying reasoning patterns (Wang et al., 2025; Leng et al., 2024). *This motivates us to further incorporate diverse perspectives into verdict reasoning, supporting perspective-aware interpretation of key cues and comprehensive reasoning for final verdicts.*

To address these, we propose *PrismJudge*, a Perspective-Conditioned Multimodal Reasoning Framework for EDV, as depicted in Fig. 1. Like a prism that reveals different perspectives of the same object, *PrismJudge* models perspective salience as learnable conditions, which jointly guide visual cue grounding and verdict decisions. Specifically, for each condition, Perspective-Conditioned Visual Cue Grounding (PCVCG) scores informative visual tokens to construct perspective-specific sparse visual representations, providing complementary visual cue supports for subsequent reasoning. Building upon these perspective-specific cues, Perspective-Conditioned Verdict Rationale Generation (PCVRG) further injects a case-adaptive perspective state into the reasoning process. Meanwhile, it introduces an asymmetric perspective-conditioned LoRA adapter to generate verdicts and rationales, where the shared A captures common dispute regulations while branch-specific B_k learns perspective-specific verdict patterns, supporting the participation of perspective conditions in interpreting key cues and providing comprehensive reasoning for final verdicts. Finally, to encourage different branches to develop complementary verdict paths, Rationale-Aligned Joint Learning (RAJL) aligns perspective branches with human rationales and jointly optimizes generation, alignment, and diversity objectives under a consistent verdict objective.

Extensive experiments demonstrate the effectiveness of perspective-conditioned cue grounding and verdict reasoning in *PrismJudge*. Compared with the strongest baseline *CyberJurors*, *PrismJudge* improves overall accuracy by 8.58%, with gains of 13.12% and 5.88% on buyer-win and seller-win cases, respectively. It also reduces the buyer–seller accuracy gap by 7.24%, indicating more accurate and balanced dispute verdicts. Furthermore, *PrismJudge* surpasses direct LoRA fine-tuning by 7.68% accuracy, highlighting the benefit of explicitly modeling perspective-aware reasoning beyond parameter-efficient adaptation.

To summarize, our main contributions are as follows:

- We propose *PrismJudge*, a perspective-conditioned multimodal reasoning framework for EDV. By learning diverse perspective conditions with RAJL, *PrismJudge* jointly modulates visual cues grounding and verdict reasoning, achieving more accurate, balanced, and reliable dispute verdicts.
- We propose PCVCG, which exploits perspective-conditioned visual scores to derive sparse visual representations, fostering complementary visual cue grounding from redundant evidence.
- We design PCVRG, which incorporates case-adaptive perspective states and perspective-conditioned LoRA into the reasoning process, sharing common verdict regulations while capturing perspective-specific reasoning patterns for cue interpretation and responsibility attribution.

2 Related Work

Dispute Verdict. Recently, MAS have excelled in solving complex problems from multiple perspectives(Zhou et al., 2026; Hong et al., 2024; Yuan & Xie, 2025; Sun et al., 2025), providing fundamental support for intelligent dispute verdicts. These methods simulate interactions among agents with diverse role prompts, aiming to encourage divergent thinking and provide more comprehensive decision support (Liang et al., 2024). For example, *AgentCourt* simulates adversarial trials with judges and opposing lawyers (Chen et al., 2025), while *TED* introduces structured debates among agents with conflicting stances (Liu et al., 2025). *CyberJurors* models jurors with diverse backgrounds to aggregate independent analyses into collective verdicts (Sun et al., 2026). However, existing empirical studies and analyses suggest that the effectiveness of role-based MAS approaches remains constrained by the inherent biases of the underlying models (Taubenfeld et al., 2024).

Conditional Representation Learning. Crowdsourced jury decisions contain fine-grained multi-perspective rationales beyond final verdicts, as diverse jurors emphasize different facts to form distinct responsibility attribution paths. (Gordon et al., 2022; Vitsakis et al., 2024). Therefore, learning latent verdict perspectives from these rationales provides a promising direction for mitigating inherent biases in verdict prediction. Conditional representation learning explores how external conditions can modulate visual representations and model adaptation, providing an effective mechanism for incorporating perspective information into verdict processes (Chen et al., 2024; Hu et al., 2022; Zhang et al., 2025). For example, InstructBLIP extracts task-relevant visual representations through an instruction-conditioned Q-Former (Dai et al., 2023), while HydraLoRA enables conditional parameter adaptation through low-rank updates (Tian et al., 2024). However, existing methods typically derive conditions from external instructions, task attributes, or predefined labels, leaving unexplored how latent perspectives can be learned from real-world rationales and propagated to visual evidence grounding and verdict generation.

3 Method

Formally, a dispute case $\mathcal{C} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_E\}$ consists of an ordered sequence of multimodal evidence items. Each item is represented as $\boldsymbol{e}_e = \{T_e, V_e^{\mathrm{raw}}\}$, containing textual claim T_e and visual information V_e^{raw}. A frozen encoder maps $\boldsymbol{e}_e$ into a text representation t_e and a set of visual tokens $\mathcal{V}_e = \{v_{e,j}\}_{j=1}^{J_e}$, where J_e denotes the number of visual tokens. Meanwhile, each case is associated with a ground-truth verdict $y_{\mathcal{C}}^{\mathrm{GT}}$ and an unordered set of human opinions $\{(y_j^{\mathrm{human}}, R_j^{\mathrm{human}})\}$, where y_j^{human} and R_j^{human} denote the verdict and rationale of the j-th human juror, respectively.

Framework Overview. To achieve perspective-aware evidence grounding and verdict generation in EDV, we propose *PrismJudge*, a perspective-conditioned multimodal reasoning framework. As illustrated in Fig. 2, *PrismJudge* consists of three components: 1) Perspective-Conditioned Visual Cue Grounding (PCVCG) leverages Perspective-Conditioned Visual Score to steer different branches toward distinct evidence regions under the same case context. Through a Sparse Visual Selection mechanism, PCVCG preserves informative visual tokens to construct perspective-specific sparse

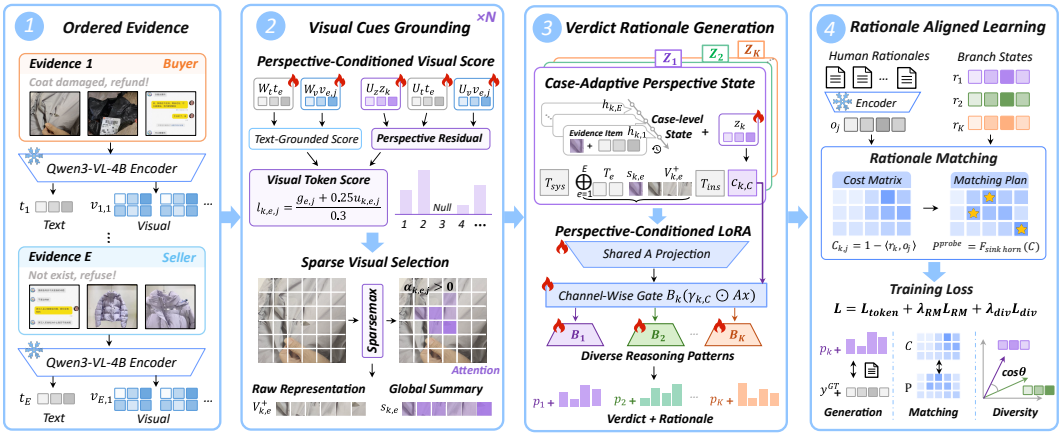


Figure 2: Overview of *PrismJudge*. Given dispute evidence, it first performs visual cue grounding to extract branch-specific visual supports, then incorporates perspective conditions into verdict rationale generation through case-adaptive states and perspective-conditioned LoRA. Finally, it applies rationale aligned learning to achieve complementary reasoning across branches.

visual representations. Building upon this, 2) Perspective-Conditioned Verdict Rationale Generation (PCVRG) first constructs a Case-Adaptive Perspective State and injects it into the model prefix $X_{k,\mathcal{C}}^{\text{prefix}}$, allowing perspective conditions to participate in the verdict reasoning process. It further introduces a Perspective-Conditioned LoRA Adapter to capture shared verdict regulations and perspective-specific reasoning patterns, and to generate responses in a "verdict–rationale" order. 3) Rationale-Aligned Joint Learning (RAJL) establishes semantic correspondences between perspective branches and human rationales through Rationale Matching, and jointly optimizes generation, rationale alignment, and branch diversity objectives to encourage complementary reasoning trajectories under a unified verdict objective. More algorithmic details are provided in the Appendix.

3.1 Perspective-Conditioned Visual Cue Grounding

Disputants often submit redundant multimodal evidence to support their respective claims, yet the key cues that clarify the dispute focus are often subtle. This requires models to ground dispute-relevant visual cues amid redundant evidence (Sun et al., 2026). Meanwhile, individuals may attend to different visual cues due to diverse visual salience (Simon, 2004; Wu & Xie, 2024). Inspired by these observations, we propose the Perspective-Conditioned Visual Cue Grounding (**PCVCG**) module. Through the perspective-conditioned visual score, PCVCG identifies cues relevant to each perspective within the same case context and constructs perspective-specific sparse visual representations, providing distinct key visual cues for subsequent verdict rationale generation.

Perspective-Conditioned Visual Score. To extract key visual cues from redundant evidence, PCVCG introduces a Perspective-Conditioned Visual Score that accounts for both evidence relevance and diverse visual salience. The score combines a text-grounded score $g_{e,j}$ with a perspective-specific residual score $u_{k,e,j}$. Specifically, the former $g_{e,j}$ computes the cosine similarity between the projected text t_e and visual representations $v_{e,j}$ to identify fine-grained visual cues aligned with the dispute claims. Furthermore, the latter $u_{k,e,j}$ first combines the projected perspective condition z_k and text representation t_e through element-wise multiplication $\odot$. Then, it measures the similarity between the resulting interaction and the projected visual representation $v_{e,j}$ to learn perspective-specific visual attention adjustments within the same evidence context. Finally, the centered and scale-calibrated residual $\widehat{u}_{k,e,j}$ is combined with the text-grounded score $g_{e,j}$ to obtain $l_{k,e,j}$:

$$g_{e,j} = \langle W_t t_e, W_v v_{e,j} \rangle,\ u_{k,e,j} = \langle U_z z_k \odot U_t t_e, U_v v_{e,j} \rangle,\ l_{k,e,j} = \frac{g_{e,j} + \lambda_v \widehat{u}_{k,e,j}}{\tau_v}. \tag{1}$$

Here, $\langle \cdot, \cdot \rangle$ denotes the inner product of two ℓ_2-normalized vectors. The λ_v controls the contribution of the perspective-conditioned residual score $\widehat{u}_{k,e,j}$relative to the text-grounded score $g_{e,j}$, while the temperature τ_v adjusts the concentration of the visual grounding score $l_{k,e,j}$.

Sparse Visual Selection. To emphasize informative visual cues while reducing redundant inputs, PCVCG applies a sparse selection mechanism. This mechanism feeds the visual scores $l_{k,e,j}$ into Sparsemax (Martins & Astudillo, 2016) to obtain sparse selection weights $\alpha_{k,e,j}$, adaptively deter-

mining the visual tokens retained from each evidence item e_e. The indices of visual tokens with positive weights are then included in the support set $\mathcal{S}_{k,e}$:

$$\{\alpha_{k,e,j}\}_{j=1}^{J_e+1} = \mathcal{F}_{Sparsemax}([l_{k,e,1}, \ldots, l_{k,e,J_e}, b_\varnothing]) \Rightarrow \mathcal{S}_{k,e} = \{j : \alpha_{k,e,j} > 0,\ j \le J_e\}, \tag{2}$$

where the learnable null score $b_\varnothing$ allows the model to select no visual token when e_e contains no informative visual cues, *i.e.,* $\mathcal{S}_{k,e} = \varnothing$, avoiding the forced retention of irrelevant information. Based on the selected tokens, PCVCG preserves their visual representations $\mathcal{V}_{k,e}^{+}$ and further compresses the attended cues into a global summary $s_{k,e}$. This design allows the verdict model to capture overall evidence patterns while explicitly reading specific supporting cues:

$$\mathcal{V}_{k,e}^{+} = \{W_s v_{e,j} : j \in \mathcal{S}_{k,e}\},\ s_{k,e} = W_s \sum\nolimits_{j=1}^{J_e} \alpha_{k,e,j} v_{e,j}, \tag{3}$$

where the shared projection W_s maps the selected visual representations and their weighted sum into the verdict space, producing $\mathcal{V}_{k,e}^{+}$ and $s_{k,e}$, respectively.

3.2 Perspective-Conditioned Verdict Rationale Generation

PCVCG provides perspective-relevant visual cues for E-commerce dispute verdicts. However, merely modifying visual inputs may not be sufficient to alter the underlying reasoning process (Wang et al., 2025), failing to induce perspective-specific reasoning patterns (Pennington & Hastie, 1986; Kahan et al., 2008). Therefore, we propose Perspective-Conditioned Verdict Rationale Generation (PCVRG), which further injects the same perspective condition as PCVCG into cues interpretation and responsibility reasoning. Specifically, PCVRG first integrates perspective conditions with case evidence to construct a case-adaptive perspective state $c_{k,\mathcal{C}}$, which serves as a condition token for the verdict model. It then employs a perspective-conditioned low-rank adapter to derive layer-wise adaptive parameter updates, jointly guiding verdict prediction p_k and rationale generation $\widehat{R}_{k,\mathcal{C}}$.

Case-Adaptive Perspective State Input. To enable perspective conditions to participate in verdict reasoning with case context, PCVRG first integrates perspective representations with case evidence to construct a case-adaptive condition state, which is then injected into the model as a condition token. Specifically, the text representation t_e of each evidence item is fused with the perspective-specific visual summary $s_{k,e}$ to obtain the evidence-level representation $h_{k,e} = t_e + \mathcal{F}_{\mathrm{L2}}(s_{k,e})$. The representations $h_{k,e}$ are aggregated and fused with the perspective condition z_k to construct the case-adaptive perspective state. Furthermore, PCVRG constructs the input sequence $X_{k,\mathcal{C}}^{\mathrm{prefix}}$ by concatenating the system prompt T_{sys}, case context, verdict instruction $T_{\mathrm{instruction}}$, and the case-adaptive perspective state $c_{k,\mathcal{C}}$:

$$c_{k,\mathcal{C}} = W_c\Big(W_q(\frac{1}{E}\sum_{e} h_{k,e}) + z_k\Big),\ X_{k,\mathcal{C}}^{\mathrm{prefix}} = \big[T_{\mathrm{sys}}; \bigoplus\nolimits_{e=1}^{E}(T_e; s_{k,e}; \mathcal{V}_{k,e}^{+}); T_{\mathrm{instruction}}; c_{k,\mathcal{C}}\big], \tag{4}$$

where the concatenation operator $\bigoplus$ follows the original order of evidence submissions, and each item is composed of its text tokens T_e, visual summary $s_{k,e}$, and selected visual tokens $\mathcal{V}_{k,e}^{+}$.

Perspective-Conditioned Low-Rank Adapter. To incorporate perspective conditions into subsequent verdict reasoning, we propose a Perspective-Conditioned Low-Rank Adapter. It incorporates $c_{k,\mathcal{C}}$ into layer-wise parameter adaptation, allowing the perspective used for visual grounding to further guide evidence interpretation and responsibility reasoning. Since different perspectives share common verdict regulations while requiring perspective-specific reasoning patterns, we adopt an asymmetric LoRA structure. Specifically, for the l-th layer, the input projection $A^{(l)}$ first maps the hidden representation into a low-rank space. After that, a learnable gating matrix $G^{(l)}$ generates channel-wise coefficients $\gamma_{k,\mathcal{C}}^{(l)}$ from $c_{k,\mathcal{C}}$, dynamically adjusting the contribution of each low-rank channel according to the current perspective condition while sharing the gating mechanism across branches. Built on this, these condition-modulated features are subsequently projected back to the original output space through the branch-specific projection $B_k^{(l)}$, enabling each perspective to learn distinct adaptation directions for perspective-aware rationale generation as follows:

$$\gamma_{k,\mathcal{C}}^{(l)} = \mathbf{1} + a_g \tanh(G^{(l)} c_{k,\mathcal{C}}),\ f_{k,\mathcal{C}}^{(l)}(x^{(l)}) = W_0^{(l)} x^{(l)} + \frac{\alpha_{\mathrm{L}}}{r_{\mathrm{L}}} B_k^{(l)}[\gamma_{k,\mathcal{C}}^{(l)} \odot A^{(l)} x^{(l)}], \tag{5}$$

where $W_0^{(l)}$ is the frozen pretrained weights, and a_g controls the magnitude of the condition gate.

Based on the perspective-conditioned LoRA, PCVRG adopts a "verdict–rationale" autoregressive generation paradigm. Specifically, each branch first outputs a verdict token to obtain verdict deci-

sions with the probability distribution p_k. Then, it continues autoregressive generation conditioned on the predicted verdict token to produce the corresponding rationale $\widehat{R}_{k,\mathcal{C}}$. This generation order prevents rationale-induced label leakage during training, as formulated below:

$$p_k, \widehat{R}_{k,\mathcal{C}} = \mathcal{F}_{k,\theta}\Big(X_{k,\mathcal{C}}^{\text{prefix}}; \{f_{k,\mathcal{C}}^{(l)}\}_l\Big). \tag{6}$$

3.3 Rationale-Aligned Joint Learning

Once the perspective-conditioned verdict model is established, a key challenge is how to further encourage different branches to develop diverse evidence attention and verdict patterns. In EDV scenarios, human jurors often refer to different facts to construct different argumentation paths, providing fine-grained supervision beyond category labels (Camburu et al., 2018). Inspired by the matching mechanism in DETR (Carion et al., 2020), we propose Rationale-Aligned Joint Learning (RAJL). It first establishes semantic correspondences between perspective branches and human rationales. Using the matching rationales, RAJL then jointly optimizes generation, alignment, and diversity objectives to learn complementary verdict paths across different branches.

Rationale Alignment. To leverage human rationales for supervising perspective-specific verdict reasoning, we first compute the semantic matching cost $C_{k,j}$ between each branch state and candidate rationales. Specifically, we extract the pre-generation state representation r_k of the k-th branch. The candidate set $\mathcal{J}_\mathcal{C}$ contains the indices of human rationales whose associated verdicts match $y_\mathcal{C}^{\text{GT}}$. The cost $C_{k,j}$ measures the semantic inconsistency between r_k and o_j using cosine distance. Since the rationale set is unordered and its size may differ from the branch count K, we perform Sinkhorn matching (Cuturi, 2013) on the cost matrix C to obtain a matching solution P^{probe}. Based on this, we perform global rationale assignment for branch k with the target rationale index $j_k^\star$ and construct the corresponding generation supervision sequence $y_k^\star$:

$$C_{k,j} = 1 - \langle r_k, o_j \rangle \xrightarrow{\mathcal{F}_{\text{Sinkhorn}}} P^{\text{probe}} \xrightarrow{\text{Assign:}j_k^\star} y_k^\star = \big[y_\mathcal{C}^{\text{GT}}; R_{j_k^\star}^{\text{human}}\big]. \tag{7}$$

Under marginal constraints, the matching solution globally aligns branches with candidate rationales while avoiding concentration on a few rationales. More details are provided in the Appendix.

Joint Optimization. Based on the above rationale alignment, RAJL jointly optimizes three objectives: autoregressive generation, rationale semantic alignment, and branch diversity. Specifically, L_{token} supervises complete verdict–rationale generation, L_{RM} aligns pre-generation branch states with human rationale representations, and L_{div} encourages different branches (k, l) to maintain diverse representations and evidence attention. The overall training objective is as follows:

$$\begin{aligned} L &= L_{\text{token}} + \lambda_{\text{RM}} L_{\text{RM}} + \lambda_{\text{div}} L_{\text{div}} \\ &= -\frac{1}{K}\sum\nolimits_{k=1}^{K}\sum\nolimits_{t=1}^{T_k} \log p_{\mathcal{F}_{k,\theta}}\left(y_{k,t}^\star \mid X_{k,\mathcal{C}}^{\text{prefix}}, y_{k,<t}^\star\right) + \lambda_{\text{RM}}\sum\nolimits_k\sum\nolimits_{j\in\mathcal{J}_\mathcal{C}} P_{k,j}^{\text{probe}} C_{k,j} \\ &+ \frac{\lambda_{\text{div}}}{|\mathcal{P}|}\sum\nolimits_{(k,l)\in\mathcal{P}}\Big[\langle z_k, z_l\rangle^2 + \lambda_1\left[m_{\text{att}} - \tfrac{1}{2}\|\bar{a}_k - \bar{a}_l\|_1\right]_+^2 + \lambda_2\left[< c_{k,\mathcal{C}}, c_{l,\mathcal{C}} > - m_{\text{cog}}\right]_+^2\Big], \end{aligned} \tag{8}$$

where $p_{\mathcal{F}_{k,\theta}}(\cdot|\cdot)$ denotes the autoregressive token distribution, and T_k denotes the effective supervision length. λ_{RM} and λ_{div} control the weights of rationale alignment and branch diversity objectives.

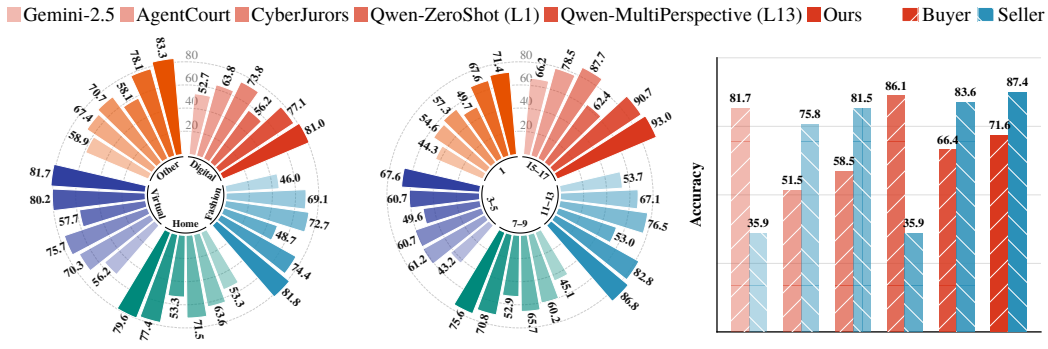
4 Experiments

Configurations. *PrismJudge* uses Qwen3-VL-4B (Team, 2025c) as the backbone. Normalization layers are applied during feature fusion to align the representation scales of heterogeneous features. Video evidence is uniformly sampled by duration, yielding up to 8 key frames per video. The LoRA rank and scaling factor are set to 16 and 32, respectively. $\lambda_v = 0.25$, $\tau_v = 0.3$, and $a_g = 0.5$. The maximum sequence length is set to 8,192 tokens. *PrismJudge* is trained for 3 epochs, with the checkpoint yielding the highest validation accuracy selected for evaluation on the full test set. By default, we report results using the equally weighted average of verdict probabilities across all K branches. We use AdamW for optimization, with peak learning rate of 2×10^{-4} for the trainable model parameters. The learning rate is linearly warmed up over the first 5% of training steps and then decayed to zero using a cosine schedule.

Evaluation. We evaluate the effectiveness of *PrismJudge* against three types of baselines: *1) LLMs* (DeepSeek-AI, 2026; OpenAI, 2026; Zeng et al., 2024), *2) MLLMs* (OpenAI, 2026; Team, 2025a;c), and *3) Dispute Verdict MAS* (Chen et al., 2025; Liu et al., 2025; Sun et al., 2026). We use the same system prompt and detailed verdict-rule instructions for *PrismJudge*, all ablation variants,

Table 1: Performance comparisons against three types of baselines. **Best** and Second Best results are highlighted. The "↓"/"↑" indicates that lower/higher values are better.

Type	Method	Acc ↑	Weig. F1 ↑	F1 ↑	Recall ↑	Precision ↑	Buyer ↑	Seller ↑
LLMs	DeepSeek-V4-Flash	0.6569	0.6603	0.6560	0.6890	0.6812	0.8153	0.5628
	Dolphin3.0-R1-24B	0.4948	0.4568	0.4790	0.5747	0.6065	0.8883	0.2612
	GPT-5.5	0.6633	0.6676	0.6618	0.6912	0.6815	0.8005	0.5819
	GLM-4	0.5017	0.4655	0.4873	0.5823	0.6180	**0.8983**	0.2663
MLLMs	Gemini-2.5-Flash-Lite	0.5292	0.5166	0.5262	0.5876	0.5987	0.8166	0.3585
	Gemini-3.1-Flash-Lite	0.6429	0.6470	0.6417	0.6719	0.6640	0.7858	0.5580
	Claude-Opus-4.5	0.5910	0.5901	0.5910	0.6344	0.6357	0.8046	0.4642
	GPT-5.5	0.6589	0.6616	0.6584	0.6947	0.6883	0.8353	0.5541
	Grok-4	0.5436	0.5512	0.5363	0.5471	0.5441	0.5609	0.5334
	Llama-3.2-90B-Vision	0.4090	0.4187	0.3987	0.4018	0.4081	0.3735	0.4300
	Qwen3-VL-4B	0.5456	0.5315	0.5429	0.6096	0.6309	0.8608	0.3585
MAS	AgentCourt	0.6673	0.6644	0.6383	0.6365	0.6413	0.5154	0.7576
	TED	0.6683	0.6587	0.6269	0.6239	0.6387	0.4498	0.7981
	CyberJurors	0.7292	0.7258	0.7037	0.6999	0.7100	0.5850	0.8148
Ours	Qwen3-VL-4B	**0.8150**	**0.8135**	**0.7991**	**0.7949**	**0.8046**	0.7162	**0.8736**



(a) Product Categories (b) Difficulty Levels (c) Accuracy Gap

Figure 3: Accuracy analysis of six models across product, difficulty, and win categories.

and all comparison baselines, as detailed in Appendix A.7. We conduct experiments on the first open-source benchmark for EDV namely *VerdictBench* (Sun et al., 2026), following the same data-splitting protocol. Meanwhile, we adopt Accuracy (Acc.), Buyer-win Accuracy (Buyer), Seller-win Accuracy (Seller), Weighted F1 (Weig. F1), Macro F1 (F1), Macro Recall, and Macro Precision (Hossin & Sulaiman, 2015) as metrics. Additionally, we compute the Accuracy Gap (Zhang & Long, 2021) and Worst-Group Accuracy (WGA) (Sagawa et al., 2019) to assess verdict bias across buyer-win and seller-win cases.

4.1 Comparisons with State-of-the-Art

Overall Performance. To evaluate the efficacy, we conduct comparative experiments on *VerdictBench*, as shown in Table 1. With the Qwen3-VL-4B backbone, *PrismJudge* achieves the best overall performance among all compared baselines. Specifically, *PrismJudge* surpasses the strongest baseline, *CyberJurors*, by 8.58 % Accuracy. It also exceeds the strongest LLM and MLLM baselines, by 15.17% and 15.61% Accuracy, respectively. The gains are consistent with the joint contribution of perspective-conditioned visual cues grounding, perspective-conditioned verdict rationale generation, and multi-perspective rationale alignment. Moreover, using the stronger *Qwen3.5-4B* (Qwen Team, 2026) backbone increases accuracy from 81.50% to 82.44%. This indicates that its gains arise from modeling the multi perspective reasoning, rather than merely compensating for the weaker backbone. More details about cross-model validations are provided in the Appendix.

Product & Difficulty. To evaluate the generalizability of *PrismJudge*, we compare it with *Gemini-2.5-Flash-Lite*, *AgentCourt*, and *CyberJurors* across diverse product categories and difficulty levels, as shown in Fig. 3(a-b). Specifically, *PrismJudge* achieves strong category robustness, with only a 3.7% accuracy variation across product categories, compared with 5.0% for the second-best baseline *CyberJurors*. It also maintains robust performance across different difficulty levels, demonstrating

Table 2: Ablation Analysis on Each Component.

ID	Variant	K	Acc.↑	WGA↑	Weig. F1↑	F1↑
Perspective-Conditioned Cues Grounding and Verdict Generation						
L1	Qwen3-VL-4B Zero-shot	1	54.56	35.85	53.15	54.29
L2	+ Opinion LoRA SFT	1	73.82	56.89	73.27	70.95
L3	+ Text-grounded Score $g_{e,j}$	1	75.56	57.30	74.90	72.60
L4	+ PCVCG	1	77.06	67.60	76.98	75.30
L5	+ PCVRG	1	78.35	68.81	78.25	76.64
Shared versus Perspective-Specific LoRA under Multi-Perspective Settings						
L6	+ Shared A, B	3	<u>80.15</u>	57.83	<u>79.16</u>	76.99
L7	**+ Branch-specific B_k (Ours⋆)**	3	**81.50**	<u>71.62</u>	**81.35**	**79.91**
L8	Independent z_k	3	79.00	67.60	78.80	<u>77.13</u>
Training Objective Controls						
L9	w/o Diversity Loss	3	78.45	**72.16**	78.50	77.06
L10	w/o Matching Loss	3	78.60	64.93	78.27	76.45
L11	w/o Rationale Matching	3	77.81	64.26	77.48	75.62
Comparisons on Multi-Perspective Baselines without PCVCG and PCVRG						
L12	Shared A, B	3	77.66	68.94	77.61	76.00
L13	HydraLoRA	3	77.21	66.40	77.05	75.31

strong generalization to challenging cases. The robustness is attributed to the complementary visual cues grounding and decision reasoning of multiple perspective branches, which enable more comprehensive coverage of diverse verdict patterns.

Accuracy Gap. To further analyze verdict bias, we compare the accuracy gap between buyer-win and seller-win cases, as reported in Table 1 and Fig. 3(c). Existing LLMs and MLLMs exhibit substantial decision asymmetry. For example, *Gemini-2.5-Flash-Lite* achieves 81.66% accuracy on buyer-win cases but only 35.85% on seller-win cases. Although MAS-based methods introduce diverse roles and agent interactions, they still suffer from notable decision imbalance, with *CyberJurors* showing a buyer/seller gap of 22.98%. In contrast, *PrismJudge* achieves more balanced performance across the two categories, reducing the gap to 15.74%, with accuracy gains of 13.12% and 5.88% on buyer-win and seller-win cases. This demonstrates that explicitly modeling complementary perspectives can alleviate the inherent category bias of foundation models and lead to more balanced verdict patterns.

4.2 Ablation Study

To evaluate the contribution of each component, we conduct ablation studies from four aspects, with the results summarized in Table 2. Detailed configurations are provided in Appendix A.4.

Perspective-Conditioned Cue Grounding and Verdict Generation. To evaluate the fine-grained contributions of perspective conditioning, we conduct incremental ablations starting from Qwen3-VL-4B zero-shot (L1). Specifically, introducing single-perspective supervision (L2) raises the accuracy from 54.56% to 73.82%. Furthermore, using the text-grounded score $g_{e,j}$ (L3) improves the accuracy to 75.56%, highlighting the benefit of selecting visual cues relevant to case claims. Moreover, PCVCG (L4) incorporates $u_{k,e,j}$ to yield 77.06% accuracy, with WGA increasing substantially from 57.30% to 67.60%. This gain suggests that perspective conditioning helps *PrismJudge* identify more relevant visual cues for decision making. Finally, PCVRG (L5) improves the accuracy to 78.35% and increases WGA to 68.81%, supporting the role of perspective-conditioned rationale generation in translating visual cues into more accurate and fair verdicts.

Perspective-Specific LoRA adapter. Extending the single-perspective LoRA (L5) to three perspectives (L6) improves accuracy from 78.35% to 80.15%, highlighting the benefits of multi-perspective modeling. However, sharing A, B (L6) integrates knowledge from complementary perspectives but may limit the modeling of perspective-specific differences, with WGA dropping from 68.81% to 57.83%. Furthermore, introducing branch-specific B_k while keeping A shared (L7) raises accuracy and WGA to 81.50% and 71.62%, respectively. The results support our asymmetric low-rank adapter design: shared A captures common dispute regulations, while branch-specific B_k learn perspective-specific differences, improving overall and worst-group accuracy. Moreover, using independent z_k

Table 3: Performance with different K.

K	Acc.↑	Weig. F1↑	F1↑	WGA↑	Acc. Gap↓
2	80.75	80.19	78.35	63.72	27.14
3	**81.50**	**81.35**	**79.91**	**71.62**	**15.74**
5	81.40	81.19	79.68	70.28	17.72

Table 4: Category accuracy across branches.

	Digital	Fashion	Home	Virtual	Other
1	80.43	81.36	81.19	**82.58**	81.37
2	**81.78**	81.36	79.62	80.78	81.64
3	80.81	**81.57**	**81.50**	81.68	**83.56**

for PCVCG and PCVRG (L8) lowers accuracy to 79.00% and WGA to 67.60%, highlighting the importance of shared perspective conditioning across cue grounding and verdict generation.

Training Objective Controls. To assess the contribution of RAJL, we compare *PrismJudge* (L7) with three ablation variants. First, removing the diversity loss (L9) lowers accuracy from 81.50% to 78.45%, with a slight WGA increase. This result is consistent with its intended role in promoting complementary reasoning rather than directly optimizing worst-group accuracy. Second, removing the matching loss (L10) reduces accuracy to 78.60% and WGA to 64.93%, highlighting the importance of direct semantic supervision from matched human rationales. Building upon this, replacing rationale matching with random assignment (L11) lowers accuracy to 77.81%, supporting the benefit of global rationale assignment.

Comparison with Multi-Perspective Baselines. To assess whether *PrismJudge* offers benefits beyond multi-perspective baselines, we compare *PrismJudge* (L7) with shared-A, B (L12) and HydraLoRA (L13). Specifically, *PrismJudge* achieves 81.50% accuracy and 71.62% WGA, outperforming L12 and L13 in accuracy by 3.84% and 4.29%, respectively. Notably, HydraLoRA does not outperform the shared-A, B baseline, suggesting that our gains do not originate from increased parameters, but rather reasoning from the complementary perspectives. Furthermore, Fig. 3 shows consistent improvements over HydraLoRA across product categories, difficulty levels and accuracy gap, supporting the benefits of PCVCG, PCVRG, and RAJL in verdict accuracy and robustness.

4.3 Branch-Level Performance Analysis

Effect of Branch Count K. To examine the effect of K, we compare *PrismJudge* with $K = 2, 3, 5$ while keeping other settings fixed. As depicted in Table 3, $K = 3$ achieves the best performance across all metrics. With $K = 2$, WGA is 63.72% and Accuracy Gap reaches 27.14%, suggesting that few branches may limit the coverage of diverse reasoning patterns. Increasing K to 3 raises WGA to 71.62% and narrows the Accuracy Gap to 15.74%. However, increasing K further to 5 yields no additional gains, suggesting that extra branches may introduce perspective redundancy rather than useful complementary information. Therefore, we adopt $K = 3$ as the perspective count.

Branch-Level Performance across Product Categories. To examine differences among jointly trained branches, we evaluate the three branches of *PrismJudge* on the test set. As summarized in Table 4, the branches exhibit different relative strengths across product categories, with no single branch performing best in every category. The within-category accuracy spread ranges from 0.21% to 2.19%. Specifically, Branch 2 leads in Digital, Branch 1 in Virtual, and Branch 3 in Fashion, Home, and Other. Notably, Branch 3, selected based on the highest validation accuracy, achieves 81.75% test accuracy, compared with 81.50% for the equal-weight ensemble. These results support the practical value of "*multi-branch joint training with single-branch inference*".

5 Conclusions

In this work, we propose *PrismJudge*, a perspective-conditioned multimodal reasoning framework for EDV. By learning latent perspective conditions from human rationales, *PrismJudge* jointly guides visual cues grounding and verdict reasoning to capture complementary cues interpretation and diverse responsibility attribution patterns. Specifically, PCVCG enables perspective-aware selection of dispute-relevant visual cues from redundant multimodal evidence, while PCVRG incorporates perspective conditions into the reasoning process, enabling branches to share common verdict knowledge while learning distinct reasoning viewpoints for the final verdicts. Extensive experiments demonstrate that *PrismJudge* achieves substantial improvements in verdict accuracy and reduces buyer–seller accuracy gaps compared with existing approaches, highlighting the effectiveness of perspective-conditioned reasoning for reliable and balanced dispute verdicts.

AI USE STATEMENT

We use large language models (LLMs) to assist with manuscript drafting and language polishing, as well as to discuss the interpretation of experimental results, without altering the technical conclusions or empirical results. All AI-assisted content is reviewed and edited by the authors. The reported empirical results are obtained through the experiments described in the paper, and the authors assess and finalize all technical claims and conclusions based on experimental evidence. The authors take full responsibility for the accuracy, originality, and integrity of the final manuscript.

ETHICS STATEMENT

This study introduces *PrismJudge*, a perspective-conditioned multimodal reasoning framework that jointly guides visual cues grounding and verdict rationale generation to provide multi-perspective decision support for E-commerce disputes. We aim to mitigate verdict bias through this multi-perspective design, while recognizing that a reduced buyer–seller accuracy gap alone is insufficient to demonstrate that the resulting verdicts are fair. Given its intended role as a decision-support tool, we highlight the following ethical considerations. First, the *VerdictBench* dataset used in this study is anonymized to protect sensitive user information and is accessed with explicit authorization from its providers solely for academic research. Second, the learned perspective conditions are anonymous computational representations and are not intended to reproduce specific jurors' identities or psychological characteristics. Finally, as potential limitations may remain unidentified, model outputs should serve as auxiliary references and be reviewed against the original case materials, with final decisions remaining the responsibility of human jury. Therefore, practical deployment should incorporate human oversight, a clear definition of *PrismJudge*'s intended role scope, and accessible mechanisms for appeals and error correction to support transparent and accountable dispute verdicts.

REPRODUCIBILITY STATEMENT

To facilitate reproducibility, we describe the proposed framework, implementation settings, and experimental protocols in the main paper and appendices. The overall framework of *PrismJudge*, including Perspective-Conditioned Visual Cue Grounding (PCVCG), Perspective-Conditioned Verdict Rationale Generation (PCVRG), and Rationale-Aligned Joint Learning (RAJL), is described in Section 3, with additional algorithmic details provided in Appendix A.1. The rationale matching procedure and joint optimization objective of RAJL are further detailed in Appendices A.2 and A.3, respectively. The training configuration, including the backbone model, LoRA settings, sequence length, optimizer, learning rate schedule, number of training epochs, and checkpoint selection criterion, is provided in Section 4. This section also describes the dataset, data-splitting protocol, comparison baselines, evaluation metrics, and branch ensembling strategy. Detailed ablation configurations, including the components used in each variant, parameter sharing schemes, training objectives, and rationale assignment strategies, are provided in Appendix A.4. Additional analyses of branch-wise reasoning differences and model generalization performance are reported in Appendices A.5 and A.6, respectively. The prompt template used to construct model inputs is provided in Appendix A.7. Furthermore, we will release the source code, model configurations, and training scripts upon acceptance to enable independent verification and facilitate future research on perspective-conditioned multimodal reasoning.

REFERENCES

Oana-Maria Camburu, Tim Rocktäschel, Thomas Lukasiewicz, and Phil Blunsom. e-snli: Natural language inference with natural language explanations. In *Advances in Neural Information Processing Systems*, pp. 9560–9572, 2018.

Nicolas Carion, Francisco Massa, Gabriel Synnaeve, Nicolas Usunier, Alexander Kirillov, and Sergey Zagoruyko. End-to-end object detection with transformers. In *ECCV*, pp. 213–229, 2020.

Chi-Min Chan, Weize Chen, Yusheng Su, et al. ChatEval: Towards better llm-based evaluators through multi-agent debate. In *The Twelfth International Conference on Learning Representations*, pp. 8190–8204, 2024.

Guhong Chen, Liyang Fan, Zihan Gong, et al. AgentCourt: Simulating court with adversarial evolvable lawyer agents. In *Findings of the Association for Computational Linguistics*, pp. 5850–5865, 2025.

Liang Chen, Haozhe Zhao, Tianyu Liu, Shuai Bai, Junyang Lin, Chang Zhou, and Baobao Chang. An image is worth 1/2 tokens after layer 2: Plug-and-play inference acceleration for large vision-language models. In *ECCV*, volume 15139, pp. 19–35, 2024.

Liyi Chen, Shuaipeng Liu, Hailei Yan, et al. ECOD: A multi-modal dataset for intelligent adjudication of e-commerce order disputes. In *CCF International Conference on Natural Language Processing and Chinese Computing*, pp. 456–468, 2023.

Marco Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in Neural Information Processing Systems*, pp. 2292–2300, 2013.

Wenliang Dai, Junnan Li, Dongxu Li, Anthony Meng Huat Tiong, Junqi Zhao, Weisheng Wang, Boyang Li, Pascale Fung, and Steven C. H. Hoi. InstructBLIP: Towards general-purpose vision-language models with instruction tuning. In *Neural Information Processing Systems*, pp. 49250–49267, 2023.

Benjamin De Haas, Alexios L Iakovidis, D Samuel Schwarzkopf, and Karl R Gegenfurtner. Individual differences in visual salience vary along semantic dimensions. *Proceedings of the National Academy of Sciences*, 116(24):11687–11692, 2019.

DeepSeek-AI. DeepSeek-V4: Towards highly efficient million-token context intelligence. *CoRR*, abs/2606.19348:1–58, 2026.

Yilun Du, Shuang Li, Antonio Torralba, et al. Improving factuality and reasoning in language models through multiagent debate. In *Proceedings of the 41st International Conference on Machine Learning*, pp. 11733–11763, 2024.

Mitchell L. Gordon, Michelle S. Lam, Joon Sung Park, Kayur Patel, Jeffrey T. Hancock, Tatsunori Hashimoto, and Michael S. Bernstein. Jury learning: Integrating dissenting voices into machine learning models. In *CHI Conference on Human Factors in Computing Systems*, pp. 115:1–115:19, 2022.

Zhitao He, Pengfei Cao, Chenhao Wang, et al. AgentsCourt: Building judicial decision-making agents with court debate simulation and legal knowledge augmentation. In *Findings of the Association for Computational Linguistics*, pp. 9399–9416, 2024.

Sirui Hong, Mingchen Zhuge, Jonathan Chen, et al. MetaGPT: Meta programming for A multi-agent collaborative framework. In *The Twelfth International Conference on Learning Representations*, pp. 35908–36008, 2024.

Mohammad Hossin and Md Nasir Sulaiman. A review on evaluation metrics for data classification evaluations. *International Journal of Data Mining & Knowledge Management Process*, 5(2):1, 2015.

Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. LoRA: Low-rank adaptation of large language models. In *The Tenth International Conference on Learning Representations*, 2022.

Zheng Jia, Shengbin Yue, Wei Chen, et al. Ready Jurist One: Benchmarking language agents for legal intelligence in dynamic environments. *CoRR*, abs/2507.04037:1–59, 2025.

Dan M Kahan, David A Hoffman, and Donald Braman. Whose eyes are you going to believe-scott v. harris and the perils of cognitive illiberalism. *Harvard Law Review*, 122:837–906, 2008.

Sicong Leng, Hang Zhang, Guanzheng Chen, Xin Li, Shijian Lu, Chunyan Miao, and Lidong Bing. Mitigating object hallucinations in large vision-language models through visual contrastive decoding. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 13872–13882, 2024.

Tian Liang, Zhiwei He, Wenxiang Jiao, et al. Encouraging divergent thinking in large language models through multi-agent debate. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pp. 17889–17904, 2024.

Yuhan Liu, Yuxuan Liu, Xiaoqing Zhang, Xiuying Chen, and Rui Yan. *The Truth Becomes Clearer Through Debate!* multi-agent systems with large language models unmask fake news. In *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 504–514, 2025.

André F. T. Martins and Ramón Fernandez Astudillo. From softmax to sparsemax: A sparse model of attention and multi-label classification. In *Proceedings of the 33nd International Conference on Machine Learning*, pp. 1614–1623, 2016.

Dorothee Mischkowski, Andreas Glöckner, and Peter Lewisch. Information search, coherence effects, and their interplay in legal decision making. *Journal of Economic Psychology*, 87:102445, 2021.

OpenAI. OpenAI GPT-5 system card. *CoRR*, abs/2601.03267:1–63, 2026.

Kerui Peng, Feifei Li, Xingyu Fan, and Wenhui Que. Semantic flow regularization: Teaching llms to generate diverse yet coherent responses. *CoRR*, abs/2605.27971, 2026.

Nancy Pennington and Reid Hastie. Evidence evaluation in complex decision making. *Journal of personality and social psychology*, 51(2):242–258, 1986.

Qwen Team. Qwen3.5-4b. Hugging Face, 2026. URL `https://huggingface.co/Qwen/Qwen3.5-4B`.

Shiori Sagawa, Pang Wei Koh, Tatsunori B. Hashimoto, and Percy Liang. Distributionally robust neural networks for group shifts: On the importance of regularization for worst-case generalization. 2019.

Dan Simon. A third view of the black box: Cognitive coherence in legal decision making. *U. Chi. L. Rev.*, 71:511, 2004.

Yanhui Sun, Wu Liu, Wentao Wang, Hantao Yao, Jiebo Luo, and Yongdong Zhang. DynamiX: Large-scale dynamic social network simulator. *CoRR*, abs/2507.19929, 2025.

Yanhui Sun, Wu Liu, Haifeng Ming, Xinru Wang, Hantao Yao, and Yongdong Zhang. CyberJurors: A multi-agent simulation task for e-commerce disputes verdict. *CoRR*, abs/2605.28369, 2026.

Amir Taubenfeld, Yaniv Dover, Roi Reichart, and Ariel Goldstein. Systematic biases in LLM simulations of debates. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pp. 251–267, 2024.

Gemini Team. Gemini 2.5: Pushing the frontier with advanced reasoning, multimodality, long context, and next generation agentic capabilities. *CoRR*, abs/2507.06261:1–72, 2025a.

Gemma Team. Gemma 3 technical report. *CoRR*, abs/2503.19786, 2025b.

Qwen Team. Qwen3-vl technical report. *CoRR*, abs/2511.21631:1–42, 2025c.

Chunlin Tian, Zhan Shi, Zhijiang Guo, Li Li, and Cheng-Zhong Xu. HydraLoRA: An asymmetric lora architecture for efficient fine-tuning. In *Advances in Neural Information Processing Systems*, 2024.

David Tsurel, Michael Doron, Alexander Nus, Arnon Dagan, Ido Guy, and Dafna Shahaf. E-commerce dispute resolution prediction. In *The 29th ACM International Conference on Information and Knowledge Management*, pp. 1465–1474, 2020.

Nikolas Vitsakis, Amit Parekh, and Ioannis Konstas. Voices in a crowd: Searching for clusters of unique perspectives. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pp. 12517–12539, 2024.

Angelina Wang, Jamie Morgenstern, and John P. Dickerson. Large language models that replace human participants can harmfully misportray and flatten identity groups. *Nature Machine Intelligence*, 7(3):400–411, 2025.

Penghao Wu and Saining Xie. V*: Guided visual search as a core mechanism in multimodal llms. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 13084–13094, 2024.

Yurun Yuan and Tengyang Xie. Reinforce LLM reasoning through multi-agent reflection. In *Forty-second International Conference on Machine Learning*, pp. 1–31, 2025.

Aohan Zeng, Bin Xu, Bowen Wang, et al. ChatGLM: A family of large language models from GLM-130B to GLM-4 all tools. *CoRR*, abs/2406.12793:1–19, 2024.

Jinghui Zhang, Kaiyang Wan, Longwei Xu, Ao Li, Zongfang Liu, and Xiuying Chen. From individuals to crowds: Dual-level public response prediction in social media. In *Proceedings of the 33rd ACM International Conference on Multimedia*, pp. 5903–5912, 2025.

Yiliang Zhang and Qi Long. Assessing fairness in the presence of missing data. In *Advances in Neural Information Processing Systems*, pp. 16007–16019, 2021.

Xuan Zhou, Yanhui Sun, Hantao Yao, Allen He, Yongdong Zhang, and Wu Liu. GASim: A graph-accelerated hybrid framework for social simulation. In *Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics*, pp. 12510–12528, 2026.

A APPENDIX

In this appendix, we provide additional details of *PrismJudge*, including the algorithm (Sec. A.1), rationale matching procedure (Sec. A.2), and optimization objective of Rationale-Aligned Joint Learning (Sec. A.3). We also provide detailed ablation configurations (Sec. A.4) and supplementary results, covering branch-wise differences (Sec. A.5) and model generalization performance (Sec. A.6), together with the complete prompt set adopted in *PrismJudge* (Sec. A.7). **Notably, the source code will be made publicly available upon the paper's acceptance.**

A.1 ALGORITHM

The algorithmic workflow of *PrismJudge* is formally presented in Algorithm 1.

Algorithm 1: *PrismJudge*: Perspective-Conditioned Multimodal Reasoning for EDV

Input: Dispute case $\mathcal{C} = \{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_E\}$, each item $\boldsymbol{e}_e = \{T_e, V_e^{\text{raw}}\}$; ground-truth verdict $y_{\mathcal{C}}^{\text{GT}}$; human opinion set $\{(y_j^{\text{human}}, R_j^{\text{human}})\}$.

Output: Branch verdict distributions $\{p_k\}_{k=1}^K$ and rationales $\{\widehat{R}_{k,\mathcal{C}}\}_{k=1}^K$

`// Perspective-Conditioned Visual Cue Grounding (PCVCG)`

Encode each evidence item: $t_e, \{v_{e,j}\}_{j=1}^{J_e} \leftarrow \mathcal{F}_{\text{encoder}}(\boldsymbol{e}_e)$;

Initialize perspective conditions: $\{z_k\}_{k=1}^K$;

for $k \leftarrow 1$ **to** K **do**

  for $e \leftarrow 1$ **to** E **do**

    Compute text-grounded score $g_{e,j}$ and perspective-specific residual score $u_{k,e,j}$: $g_{e,j} \leftarrow \langle W_t t_e, W_v v_{e,j} \rangle$; $u_{k,e,j} \leftarrow \langle U_z z_k \odot U_t t_e, U_v v_{e,j} \rangle$;

    Obtain visual grounding score: $l_{k,e,j} \leftarrow \frac{g_{e,j} + \lambda_v \widehat{u}_{k,e,j}}{\tau_v}$;

    Generate sparse selection weights: $\{\alpha_{k,e,j}\}_{j=1}^{J_e+1} \leftarrow \mathcal{F}_{\text{Sparsemax}}([l_{k,e,1}, ..., l_{k,e,J_e}, b_{\varnothing}])$;

    Select visual tokens with positive weights: $\mathcal{S}_{k,e} \leftarrow \{j : \alpha_{k,e,j} > 0, j \le J_e\}$;

    Construct visual supports $\mathcal{V}_{k,e}^+$ and cues summary $s_{k,e}$ with the selected tokens: $\mathcal{V}_{k,e}^+ \leftarrow \{W_s v_{e,j} : j \in \mathcal{S}_{k,e}\}$, $s_{k,e} \leftarrow W_s \sum_{j=1}^{J_e} \alpha_{k,e,j} v_{e,j}$.

`// Perspective-Conditioned Verdict Rationale Generation (PCVRG)`

for $k \leftarrow 1$ **to** K **do**

  Fuse textual representation with perspective-specific visual summary: $h_{k,e} \leftarrow t_e + \mathcal{F}_{\text{L2}}(s_{k,e}), \quad e = 1, \ldots, E$;

  Construct case-adaptive perspective state: $c_{k,\mathcal{C}} \leftarrow W_c \left(W_q(\frac{1}{E} \sum_{e=1}^{E} h_{k,e}) + z_k \right)$;

  Build input sequence with system prompt T_{sys} and verdict instruction $T_{\text{instruction}}$: $X_{k,\mathcal{C}}^{\text{prefix}} \leftarrow [T_{\text{sys}}; \bigoplus_{e=1}^{E} (T_e; s_{k,e}; \mathcal{V}_{k,e}^+); T_{\text{instruction}}; c_{k,\mathcal{C}}]$;

  Apply perspective-conditioned low-rank adapter: $\gamma_{k,\mathcal{C}}^{(l)} \leftarrow \mathbf{1} + a_g \tanh(G^{(l)} c_{k,\mathcal{C}})$; $f_{k,\mathcal{C}}^{(l)}(x^{(l)}) \leftarrow W_0^{(l)} x^{(l)} + \frac{\alpha_L}{r_L} B_k^{(l)} [\gamma_{k,\mathcal{C}}^{(l)} \odot A^{(l)} x^{(l)}]$;

  Generate branch verdict distribution p_k and rationale $\widehat{R}_{k,\mathcal{C}}$: $p_k, \widehat{R}_{k,\mathcal{C}} \leftarrow \mathcal{F}_{k,\theta}(X_{k,\mathcal{C}}^{\text{prefix}}; \{f_{k,\mathcal{C}}^{(l)}\}_l)$.

`// Rationale-Aligned Joint Learning (RAJL)`

Extract the pre-generation state representation r_k of the k-th branch;

Encode candidate human rationales R_j^{human} from $\mathcal{J}_{\mathcal{C}}$ into semantic representation o_j;

Compute matching cost $C_{k,j} = 1 - \langle r_k, o_j \rangle$ between each branch state and candidate rationales;

Perform Sinkhorn matching on the cost matrix: $P^{\text{probe}} \leftarrow \mathcal{F}_{\text{Sinkhorn}}(C)$;

Use the global matching solution to assign a target rationale with index $j_k^\star$ to branch k;

Construct the supervision sequence: $y_k^\star = [y_{\mathcal{C}}^{\text{GT}}; R_{j_k^\star}^{\text{human}}]$;

Optimize autoregressive generation, semantic alignment and branch diversity objectives: $L = L_{\text{token}} + \lambda_{\text{RM}} L_{\text{RM}} + \lambda_{\text{div}} L_{\text{div}}$.

A.2 RATIONALE MATCHING

This section further provides implementation details of the branch–rationale matching module in RAJL. Specifically, we introduce the construction of pre-generation branch representations, candidate rationale filtering, Sinkhorn soft matching, and global rationale assignment, and illustrate how branch-level supervision targets are constructed from the matching results.

Pre-generation Representation and Reading Position. To establish semantic correspondence between perspective branches and human rationales while avoiding supervision leakage, RAJL requires branch representations that do not contain target rationale information. Therefore, we extract the continuous representation r_k from the pre-generation prefix. Specifically, we insert a shared learnable continuous vector e_{PLAN} after the case-adaptive perspective state $c_{k,C}$ and before the first verdict token:

$$\widetilde{X}_{k,C}^{\mathrm{prefix}} = [X_{k,C}^{\mathrm{prefix}}; e_{\mathrm{PLAN}}], \tag{9}$$

where $\widetilde{X}_{k,C}^{\mathrm{prefix}}$ denotes the input sequence augmented with the auxiliary reading position. For notational simplicity, we continue to use $X_{k,C}^{\mathrm{prefix}}$ to denote the autoregressive conditioning prefix below. Although e_{PLAN} is shared across different branches, its resulting hidden state $H_{k,C}^{\mathrm{PLAN}}$ is jointly determined by the input prefix, $c_{k,C}$, and the branch adapter, thereby preserving branch-conditioned information. Moreover, due to the causal attention constraint, this position cannot access the ground-truth verdict or human rationale in the target sequence. Subsequently, we project the hidden state into the rationale semantic space and normalize it to obtain the branch representation r_k:

$$r_k = \mathrm{Norm}\big(W_p\,\mathrm{LN}(H_{k,C}^{\mathrm{PLAN}})\big), \qquad \mathrm{Norm}(u) = \frac{u}{\|u\|_2}. \tag{10}$$

Here, LN denotes layer normalization. r_k characterizes the pre-generation branch state under the current case context and perspective branch, which is used for subsequent semantic matching between perspective branches and human rationales.

Candidate Rationales and Matching Cost. Based on the pre-generation branch representation r_k, we construct the candidate rationale set from the human opinion set $\{(y_j^{\mathrm{human}}, R_j^{\mathrm{human}})\}$. To prevent incorrect verdicts from interfering with the matching process, we only retain opinions whose human verdicts are consistent with the ground-truth verdict:

$$\mathcal{J}_C = \{j : y_j^{\mathrm{human}} = y_C^{\mathrm{GT}},\ R_j^{\mathrm{human}} \text{ is a valid non-empty rationale}\}, \qquad J_C = |\mathcal{J}_C|. \tag{11}$$

For each candidate rationale R_j^{human}, we employ a frozen encoder from our backbone to extract the mean representation of valid tokens and normalize it to obtain o_j. We then compute the semantic matching cost between the branch representation and the rationale representation: $C_{k,j} = 1 - \langle r_k, o_j \rangle$.

Sinkhorn Matching. If each branch independently selects the rationale with the lowest matching cost, multiple branches may collapse onto the same rationale, leaving other candidate rationales without matching weights. Therefore, we obtain the matching plan $P^{\mathrm{probe}} = F_{\mathrm{Sinkhorn}}(C)$ from the cost matrix C, subject to prescribed marginal constraints. Specifically, by setting the branch marginal $\mu_k = 1/K$ and the rationale marginal $\nu_j = 1/J_{\mathcal{C}}$, we define

$$P^{\mathrm{probe}} \in \mathcal{U}(\mu,\nu), \quad \mathcal{U}(\mu,\nu) = \Big\{P \geq 0 : \sum_{j\in\mathcal{J}_C} P_{k,j} = \mu_k,\ \sum_{k=1}^{K} P_{k,j} = \nu_j\Big\}, \tag{12}$$

where $\mathcal{U}(\mu,\nu)$ denotes the set of all non-negative $K \times J_{\mathcal{C}}$ matrices with row sums equal to μ and column sums equal to ν. The uniform marginal constraints assign each branch a fixed total matching weight of $1/K$, while restricting the aggregated matching weight received by each candidate rationale from all branches to $1/J_{\mathcal{C}}$. This produces a more balanced branch–rationale allocation and encourages different branches to learn complementary reasoning paths. Under these constraints, to further incorporate the semantic cost $C_{k,j}$ into the matching process and obtain a soft matching matrix solvable via Sinkhorn iterations, we optimize the following objective over the feasible set:

$$P^{\mathrm{probe}} = \underset{P\in\mathcal{U}(\mu,\nu)}{\arg\min} \left\{ \sum_{k=1}^{K}\sum_{j\in\mathcal{J}_C} P_{k,j}C_{k,j} + \varepsilon \sum_{k,j} P_{k,j}(\log P_{k,j} - 1) \right\}, \qquad \varepsilon > 0. \tag{13}$$

The first term encourages assignments with lower semantic costs, while the second term preserves soft matching through entropy regularization, with $\varepsilon = 0.02$ controlling the degree of softness. During optimization, Sinkhorn iterations alternately normalize the row and column weights, driving P^{probe} toward the prescribed marginal constraints.

Global Rationale Assignment. The Sinkhorn matching result P^{probe} preserves soft matching weights over multiple candidate rationales for each branch, whereas autoregressive training requires assigning a specific rationale to each branch as its supervision target. To this end, we first normalize the k-th row using the row marginal $\mu_k = 1/K$: $q_{k,j} = P^{\text{probe}}_{k,j} / \sum_{m \in \mathcal{J}_C} P^{\text{probe}}_{k,m} = K P^{\text{probe}}_{k,j}$. This gives $\sum_{j \in \mathcal{J}_C} q_{k,j} = 1$, and $q_{k,\cdot}$ can be regarded as the soft matching distribution over candidate rationales for branch k. Then, for each candidate rationale j, we compute:

$$D_{k,j} = \sum_{m \in \mathcal{J}_C} q_{k,m} \big(1 - \langle o_j, o_m \rangle\big), \tag{14}$$

where $D_{k,j}$ measures the expected semantic distance between candidate rationale j, when assigned as the supervision rationale for branch k, and the soft-matched rationale distribution of that branch. Compared with directly selecting the rationale with the largest $P^{\text{probe}}_{k,j}$, this design integrates information from multiple semantically related rationales and avoids the selection bias caused by relying on a single highest-weight match. After that, we jointly solve the global rationale assignment across all branches:

$$(j_1^\star, \ldots, j_K^\star) \in \arg\min_{\mathbf{j} \in \mathcal{A}_{K,\mathcal{C}}} \sum_{k=1}^{K} D_{k,j_k}, \qquad \mathbf{j} = (j_1, \ldots, j_K), \tag{15}$$

where $\mathcal{A}_{K,\mathcal{C}} = \{\mathbf{j} \in \mathcal{J}_C^K \mid \forall\, k \neq l,\ j_k \neq j_l\}$ constrains different branches to be assigned different candidate rationales. Finally, following Eq. (7), we construct the branch supervision target $y_k^\star = [y_C^{\text{GT}}; R^{\text{human}}_{j_k^\star}]$. Since each case contains at least K human rationales consistent with the ground-truth verdict, *i.e.,* $J_{\mathcal{C}} \geq K$, the global assignment provides each branch with a distinct human rationale as its supervision signal.

A.3 Joint Optimization of RATIONALE-ALIGNED JOINT LEARNING

This section further provides implementation details of the RAJL joint optimization process. Specifically, we first introduce the dual-forward mechanism between matching target construction and parameter optimization to prevent supervision leakage during rationale matching. We then describe the detailed computation of the generation loss, rationale matching loss, and branch diversity loss based on the fixed matching results.

Matching Forward and Training Forward. The construction of supervision targets in Eq. (7) requires obtaining branch representations and determining matching relationships without accessing the target rationales, whereas the rationale matching loss in Eq. (8) requires gradient optimization based on the current model parameters. Therefore, we decouple rationale assignment and parameter optimization into two separate forward processes. First, we perform the matching forward pass with dropout disabled and without gradient computation. Specifically, each branch only receives the case prefix and the pre-generation readout position to yield branch representation r_k^{probe}, which is independent of the target rationales. Then, we compute the matching cost through $C^{\text{probe}}_{k,j} = 1 - \langle r_k^{\text{probe}}, o_j \rangle$, and obtain the soft matching matrix P^{probe} as well as the discrete assignment result $j_k^\star$. Since no target sequence is provided during this stage, the matching process avoids information leakage from the assigned rationales. After obtaining the fixed supervision target $y_k^\star$, we restore the training mode and perform the training forward pass. Here, the model computes the autoregressive generation loss over the complete target sequence, while simultaneously re-extracting a gradient-enabled branch representation r_k^{train} and computing: $C^{\text{train}}_{k,j} = 1 - \langle r_k^{\text{train}}, o_j \rangle$. Both forward processes share the identical evidence order, visible input prefix, and branch conditions. To maintain consistency with Eq. (8), we denote $C^{\text{train}}_{k,j}$ obtained from the training forward pass simply as $C_{k,j}$ in the following. Since the discrete assignment $j_k^\star$ and the soft matching matrix P^{probe} remain fixed during the current optimization step, $\text{sg}(P^{\text{probe}})$ is used only to indicate gradient stopping and does not alter the forward value of the matching loss in Eq. (8).

Joint Objective. Based on the fixed supervision targets $y_k^\star$ and the soft matching matrix P^{probe} obtained from the matching stage, RAJL jointly optimizes the autoregressive generation loss, rationale

matching loss, and branch diversity loss following Eq. (8):

$$L = L_{\text{token}} + \lambda_{\text{RM}} L_{\text{RM}} + \lambda_{\text{div}} L_{\text{div}}. \tag{16}$$

Generation Loss. For the target sequence $y_k^\star$, we compute its autoregressive negative log-likelihood:

$$L_{\text{token}} = -\frac{1}{K} \sum_{k=1}^{K} \sum_{t=1}^{T_k} \log p_{F_{k,\theta}} \left(y_{k,t}^\star \mid X_{k,C}^{\text{prefix}}, y_{k,<t}^\star \right),$$

where $p_{F_{k,\theta}}$ denotes the autoregressive token probability distribution of the k-th branch, and T_k denotes the effective supervision length of the corresponding branch. Since $y_k^\star$ contains both the ground-truth verdict and the assigned human rationale, both are optimized through L_{token} under the same autoregressive objective.

Rationale Matching Loss. Based on the fixed soft matching matrix P^{probe} obtained from the matching forward pass, we employ it as matching weights in the training forward pass to align the current branch representation r_k^{train} with the human rationale representation o_j:

$$L_{\text{RM}} = \sum_{k=1}^{K} \sum_{j \in \mathcal{J}_C} \text{sg}(P_{k,j}^{\text{probe}}) \big[1 - \langle r_k^{\text{train}}, o_j \rangle \big].$$

Accordingly, P^{probe} serves as semantic alignment weights, applying human rationale supervision to the branch internal states before verdict generation.

Branch Diversity Loss. To encourage different branches within the same case to learn complementary representations, let $\mathcal{P} = \{(k, l) : 1 \le k < l \le K\}$ denote the set of all unordered branch pairs within the same case. Based on these branch pairs, the diversity loss is defined as:

$$L_{\text{div}} = \frac{1}{|\mathcal{P}|} \sum_{(k,l) \in \mathcal{P}} \Big(\langle z_k, z_l \rangle^2 + \lambda_1 \big[m_{\text{att}} - \tfrac{1}{2} \|\bar{a}_k - \bar{a}_l\|_1 \big]_+^2 + \lambda_2 \big[\langle c_{k,C}, c_{l,C} \rangle - m_{\text{cog}} \big]_+^2 \Big), \tag{17}$$

where $|\mathcal{P}| = K(K-1)/2$ and $[x]_+ = \max(x, 0)$. m_{att} denotes the minimum threshold for inter-branch visual attention discrepancy between branches, and m_{cog} denotes the maximum threshold for the cosine similarity between branch-conditioned states. For the diversity objective, we set $\lambda_1 = \lambda_2 = 20$, $m_{\text{att}} = 0.20$, and $m_{\text{cog}} = 0.95$.

In Eq. (17), the first term reduces the correlation between perspective representations z_k and z_l, while the latter two terms introduce penalties only when the corresponding constraints are violated. The coefficients λ_1 and λ_2 control the relative contributions of the latter two terms. In the second term, a_k is constructed by collecting the attention weights $\alpha_{k,e,j}$ at all visual positions satisfying $J_e > 1$ following the order of evidence submissions. All branches share the same position indices, with NULL excluded and unselected positions zeroed. The normalized attention vector is defined as: $\bar{a}_k = a_k / \max\{\text{sg}(\|a_k\|_1), 10^{-12}\}$.

A.4 Detailed Ablation Configurations

We detail the configurations of the ablation variants in Table 2, using the same IDs L1–L13. All variants use the Qwen3-VL-4B backbone. Table 5 summarizes their architectural components, and Table 6 specifies the training-objective controls. Here, ✓ indicates an enabled component, × indicates a disabled component, and "–" denotes a setting that is not applicable.

Architectural Components. The two PCVCG columns distinguish the text-grounded score $g_{e,j}$ from the perspective-specific residual score $u_{k,e,j}$. Thus, L3 enables text-grounded visual selection only, whereas L4 additionally introduces the perspective-specific residual. The PCVRG column indicates the case-adaptive perspective-state input and perspective-conditioned LoRA gating introduced in L5. "Single" denotes a single-branch LoRA adapter, "Shared" denotes A and B shared across branches, and "Asymmetric" denotes a shared A with branch-specific B_k. The "Shared z_k" column indicates whether PCVCG and PCVRG use the same perspective condition within each branch; it does not indicate sharing one condition across all branches. L8 uses independent conditions for the two modules. L12 and L13 exclude PCVCG and PCVRG and use shared-A, B adaptation and HydraLoRA, respectively.

Training-Objective Controls. L9 removes L_{div} while retaining rationale matching and L_{RM}. L10 removes L_{RM} but retains global rationale assignment to construct the generation targets. Building

Table 5: Component configurations of the ablation variants.

ID	Variant	K	PCVCG		PCVRG	Shared z_k	LoRA
			$g_{e,j}$	$u_{k,e,j}$			
L1	Zero-shot	1	×	×	×	–	–
L2	Opinion LoRA SFT	1	×	×	×	–	Single
L3	Text-grounded score	1	✓	×	×	–	Single
L4	PCVCG	1	✓	✓	×	–	Single
L5	PCVCG + PCVRG	1	✓	✓	✓	✓	Single
L6	Shared A, B	3	✓	✓	✓	✓	Shared
L7	***PrismJudge***	3	✓	✓	✓	✓	Asymmetric
L8	Independent z_k	3	✓	✓	✓	×	Asymmetric
L9	w/o Diversity Loss	3	✓	✓	✓	✓	Asymmetric
L10	w/o Matching Loss	3	✓	✓	✓	✓	Asymmetric
L11	w/o Rationale Matching	3	✓	✓	✓	✓	Asymmetric
L12	Shared A, B baseline	3	×	×	×	–	Shared
L13	HydraLoRA baseline	3	×	×	×	–	HydraLoRA

Table 6: Training-objective configurations for the RAJL ablations.

ID	L_{token}	L_{div}	L_{RM}	Rationale assignment
L7	✓	✓	✓	Global matching
L9	✓	×	✓	Global matching
L10	✓	✓	×	Global matching
L11	✓	✓	×	Random assignment

on L10, L11 replaces global rationale assignment with random assignment while keeping L_{RM} disabled. All four variants in Table 6 retain the complete architecture of L7 and the verdict–rationale generation loss L_{token}. For the single-branch variants ($K = 1$), the branch diversity objective is omitted.

Notably, the L12-L13 comparison provides indirect evidence that the improvement from L6 to L7 is unlikely to be driven solely by the additional branch-specific B_k parameters.

A.5 Branch-Wise Differences

Beyond the category-level comparison of verdict performance, this appendix conducts a case-level analysis of different perspective branches. Specifically, we investigate the utilization of visual evidence, the characteristics of generated rationales, and the robustness of predictions when subsets of sparse cue representation are removed.

A.5.1 Perspective Differences in Visual Attention and Verdict Reasoning

To examine whether different branches learn distinct perspective patterns, we conduct pairwise comparisons among the three branches on the test set $\mathcal{D}_{\text{test}}$. For each branch pair (k, l) and case $\mathcal{C}$, we quantify the similarity of their visual supports $\mathcal{S}^{\text{vis}}_{k,\mathcal{C}}$ and $\mathcal{S}^{\text{vis}}_{l,\mathcal{C}}$ using Jaccard similarity, and measure the similarity between their case-adaptive perspective states $c_{k,\mathcal{C}}$ and $c_{l,\mathcal{C}}$ using cosine similarity. We further compute the verdict disagreement rate between the predicted verdicts p_k and p_l.

$$J^{\text{vis}}_{k,l} = \frac{1}{n} \sum_{\mathcal{C} \in \mathcal{D}_{\text{test}}} \frac{|\mathcal{S}^{\text{vis}}_{k,\mathcal{C}} \cap \mathcal{S}^{\text{vis}}_{l,\mathcal{C}}|}{|\mathcal{S}^{\text{vis}}_{k,\mathcal{C}} \cup \mathcal{S}^{\text{vis}}_{l,\mathcal{C}}|}, \quad \rho^{\text{state}}_{k,l} = \frac{1}{n} \sum_{\mathcal{C} \in \mathcal{D}_{\text{test}}} \cos(c_{k,\mathcal{C}}, c_{l,\mathcal{C}}), \tag{18}$$

where $\mathcal{S}^{\text{vis}}_{k,\mathcal{C}}$ denotes the visual support set of branch k for case $\mathcal{C}$, defined as $\mathcal{S}^{\text{vis}}_{k,\mathcal{C}} = \{(e, j) : 1 \leq e \leq E_{\mathcal{C}}, j \in \mathcal{S}_{k,e}\}$, where e denotes the evidence index and j denotes the local position of a selected visual token within the corresponding evidence item. $n = |\mathcal{D}_{\text{test}}| = 2{,}005$.

Table 7: Pairwise branch differences in visual supports, case states, and verdicts.

Branch Pair	Visual Support	Perspective States	Verdict Disagreement
(1,2)	0.4093	0.6368	5.54%
(1,3)	0.5372	0.6291	5.44%
(2,3)	0.5106	0.6791	5.79%

As shown in Table 7, branch pairs exhibit moderate visual support overlap, with Jaccard similarities ranging from 0.4093 to 0.5372. The cosine similarities of case-adaptive perspective states range from 0.6291 to 0.6791, indicating that different branches maintain distinct representations with the same case context. Meanwhile, the verdict disagreement rates among branch pairs remain around 5–6%, suggesting that different branches capture distinct decision patterns while maintaining substantial verdict consistency.

A.5.2 Branch-Level Rationale Diversity

To analyze whether different branches develop distinct reasoning preferences, we categorize the generated rationales from the three branches on the test set according to predefined semantic themes. After removing verdict prefixes and template phrases, each rationale is assigned to the theme with the highest keyword coverage among five categories: evidence sufficiency (*e.g.*, "video", "evidence submission"), description/disclosure (*e.g.*, "concealment", "inconsistency"), quality/functionality (*e.g.*, "damage", "repair"), refund/remedy (*e.g.*, "compensation", "after-sales service"), and responsibility attribution (*e.g.*, "buyer failed to install", "free-riding"). If a rationale matches multiple themes, it is assigned to the theme with the largest number of matched keywords. Ties are resolved according to the earliest keyword occurrence in the rationale. Rationales without valid keyword matches remain unclassified.

Table 8: Themes distributions of generated rationales across three branches.

Branch	Evidence Sufficiency	Description Disclosure	Quality Functionality	Refund Remedy	Responsibility Attribution
1	17.45	27.70	6.59	4.33	**43.93**
2	16.72	**45.74**	**10.90**	4.96	21.68
3	**26.36**	42.33	5.05	**6.20**	20.05

As shown in Table 8, different branches emphasize different reasoning aspects. Specifically, Branch 1 has the highest proportion of responsibility attribution (43.93%), indicating a stronger tendency toward liability-related explanations. Branch 2 places more emphasis on description/disclosure (45.74%) and quality/functionality (10.90%), focusing more on product conditions and consistency with descriptions. Branch 3 shows higher proportions of evidence sufficiency (26.36%) and refund/remedy (6.20%), suggesting greater attention to evidence support and resolution outcomes. These results indicate that, although all branches process the same case information, they develop different rationale preferences and complementary reasoning patterns.

A.5.3 Visual Cue Contribution Analysis

To evaluate whether the visual cues selected by PCVCG contribute to verdict prediction, we conduct a visual cue removal experiment. For each case, we first obtain the branch-specific selected visual tokens $\mathcal{V}_{k,e}^{+}$ produced by PCVCG. We then randomly remove approximately 25% of the selected visual tokens and perform verdict reasoning with the remaining visual supports. Before and after removal, we keep the textual evidence and model configurations unchanged, while recomputing the case-adaptive perspective state, perspective-conditioned adaptation, and final verdict prediction using the remaining visual supports. Each branch determines the final prediction according to its probability distribution over the first verdict tokens. Each branch is evaluated based on its own selected visual tokens.

As summarized in Table 9, removing a portion of the PCVCG-selected visual tokens consistently degrades the performance of all three branches, with accuracy drops of 2.05, 2.45, and 3.05 % for

Table 9: Accuracy changes after removing selected visual tokens across three branches.

Branch	Acc. Before	Acc. After	Drop
1	81.30	79.25	2.05
2	81.15	78.70	2.45
3	81.75	78.70	3.05

Branch 1, Branch 2, and Branch 3, respectively. This consistent degradation indicates that the visual tokens selected by PCVCG contain decision-relevant cues that contribute to verdict prediction.

Notably, Branch 3, selected solely on validation accuracy, achieves 81.75% test accuracy, compared with 81.50% for the equal-weight ensemble. This result suggests that joint multi-branch training can produce a strong individual branch suitable for single-branch inference, supporting the practical value of *multi-branch joint training with single-branch inference*. We do not position inference-time multi-perspective ensembling as our main contribution.

The quantitative analysis above shows that different branches maintain distinct visual supports and reasoning states. We further provide a case study to illustrate differences in visual cue selection and verdict rationales across perspective branches. Specifically, we compare branches 1 (Blue) and 2 (Green) of *PrismJudge* on test case 3013, which concerns a dispute over a second-hand automotive tail-light. As shown in Fig. 4, the case contains three chat screenshots from evidence item $e = 4$. The blue branch selects 3, 2, and 7 visual tokens from the first, second, and third screenshots, respectively, assigning 8.3%, 13.8%, and 77.9% of its visual grounding weight within this evidence item. Several selected token positions in the third screenshot are located near the video thumbnail submitted by the buyer, indicating stronger attention to the visual evidence provided by the buyer. In contrast, the green branch selects 2, 2, and 4 visual tokens from the three screenshots, with corresponding weight shares of 16.5%, 25.2%, and 58.3%. Compared with the blue branch, the green branch allocates relatively more attention to the first two screenshots, where the seller discusses the second-hand condition of the item and signs of use. The two branches exhibit substantially different visual supports within evidence item $e = 4$, with a Jaccard overlap of only 0.053. Consistent with these distinct visual emphases, branch $k = 1$ predicts Buyer and generates a rationale focusing on "*an undisclosed defect*". Meanwhile, branch $k = 2$ predicts Seller and attributes the outcome to "*the disclosure of the used condition*" and "*the buyer raised no objection*". This case demonstrates that different perspective branches can attend to different visual cues from the same evidence and develop distinct evidence interpretations for verdict reasoning.

A.6 Model Generalization Performance

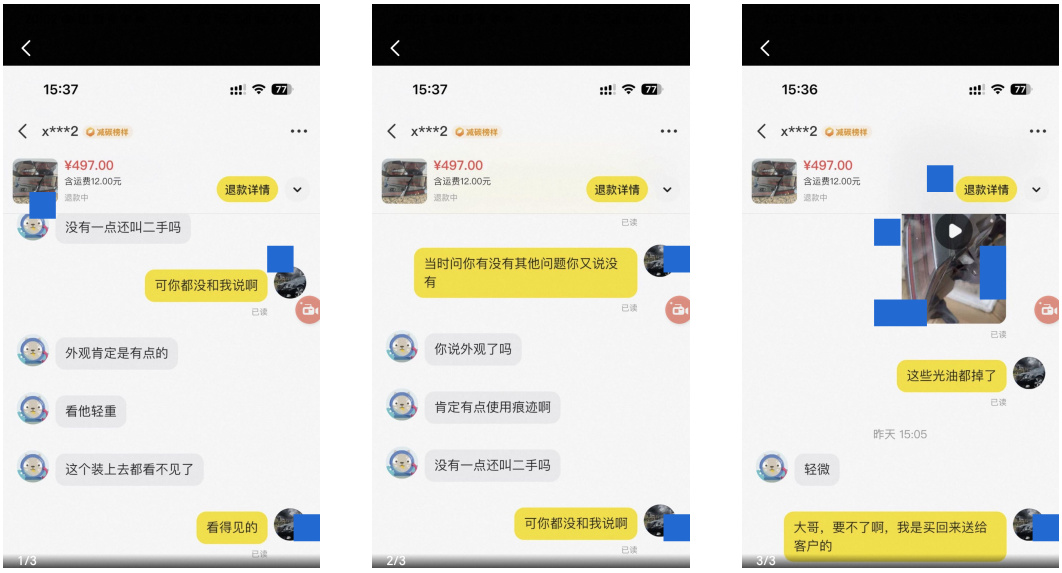
Generalization to a Stronger Backbone. To evaluate the generalizability of *PrismJudge* to a stronger backbone, we replace *Qwen3-VL-4B* with *Qwen3.5-4B* (Qwen Team, 2026). As shown in Table 10, the Qwen3.5-4B variant achieves an accuracy of 82.44%, outperforming its *Qwen3-VL-4B* counterpart by 0.94%. It also achieves higher scores on all other reported metrics. The results indicate that the gains of *PrismJudge* arise from modeling the multi perspective reasoning, rather than merely compensating for the weaker backbone.

Table 10: Performance of *PrismJudge* with different Qwen backbones.

Backbone	Acc. ↑	Weig. F1 ↑	F1 ↑	Recall ↑	Precision ↑	Buyer ↑	Seller ↑
Qwen3-VL-4B	0.8150	0.8135	0.7991	0.7949	0.8046	0.7162	0.8736
Qwen3.5-4B	**0.8244**	**0.8234**	**0.8101**	**0.8068**	**0.8142**	**0.7376**	**0.8760**

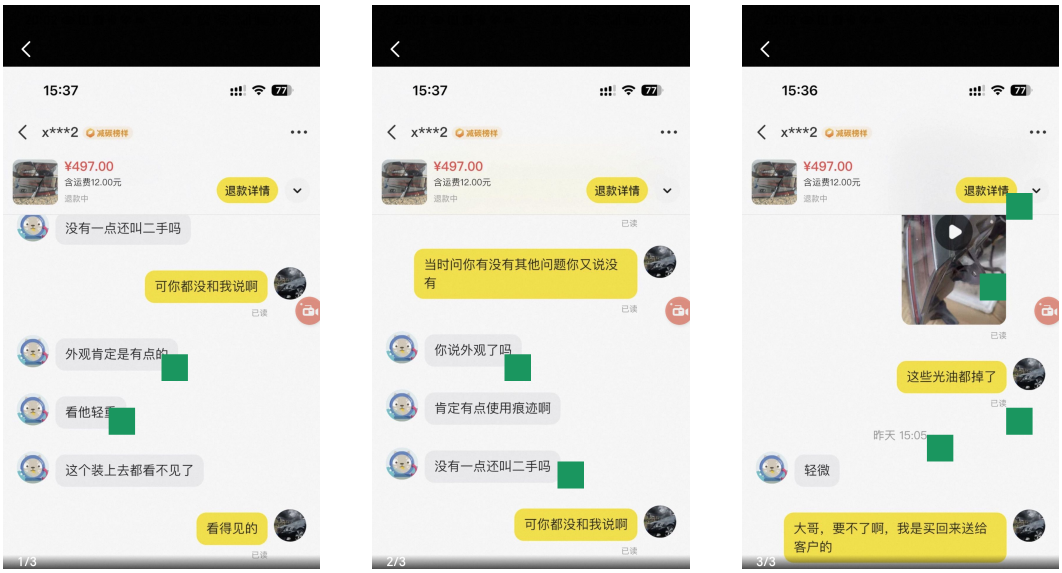
Cross-Backbone Generalization on *Gemma-3-4B-IT*. To evaluate the generalizability of *PrismJudge* beyond the Qwen backbone, we conduct cross-backbone experiments using *Gemma-3-4B-IT* (Team, 2025b) on the same test set, as detailed in Table 11. We compare *PrismJudge* with the multi-perspective HydraLoRA baseline without PCVCG and PCVRG (L13) using the same experimental settings as in the ablation study. Compared with the multi-perspective HydraLoRA baseline, *PrismJudge* improves accuracy from 75.66% to 78.50% and achieves higher Weighted F1, Macro F1, Recall, and Precision, demonstrating that the benefits of perspective-conditioned grounding and reasoning generalize across different backbones.

Verdict: Buyer Win
Rationale: The product did not function as described, and the seller concealed a defect without prior disclosure.



(a) Branch 1

Verdict: Seller Win
Rationale: No defect was concealed and the appearance matched the description. The seller had stated that the item was second-hand, and the buyer raised no objection.



(b) Branch 2

Figure 4: Case study of evidence grounding effects on verdict rationales. Colored squares (■ and ■) indicate the positions of visual tokens selected by Sparsemax.

Table 11: Cross-backbone results on *Gemma-3-4B-IT*.

Method	Acc. ↑	Weig. F1 ↑	Macro F1 ↑	Recall ↑	Precision ↑
Multi-Perspective HydraLoRA	75.66	75.53	73.70	73.48	73.98
PrismJudge	**78.50**	**77.95**	**75.96**	**75.09**	**77.81**

Performance Stability. To evaluate the robustness of *PrismJudge* to training randomness, we report its performance with the Qwen3-VL-4B backbone under three different seeds (42, 44, and 46). As shown in Table 12, *PrismJudge* maintains consistent performance across different runs. The standard deviations of all metrics remain below 0.5 percentage points, indicating that the performance of *PrismJudge* is stable across different seeds.

Table 12: Performance of *PrismJudge* under three seeds.

Seed	Acc.↑	Weig. F1↑	F1↑	Recall↑	Precision↑
42	81.50	81.35	79.91	79.49	80.46
44	81.70	81.64	80.30	80.11	80.51
46	81.70	81.68	80.40	80.36	80.44
Mean±Std.	81.63±0.12	81.56±0.18	80.20±0.26	79.99±0.45	80.47±0.04

A.7 Prompt Set

We present the prompt template used by *PrismJudge* for multimodal dispute verdict reasoning. The template includes case information, buyer and seller claims, and multimodal evidence arranged in the original submission order, together with verdict criteria and output requirements. It instructs the model to first output a verdict token, `buyer` or `seller`, followed by a rationale that explicitly cites the decisive evidence. The template is shared across perspective branches, with each branch receiving its corresponding visual cues and perspective condition.

Prompt A.7.1: Perspective-Conditioned Multimodal Reasoning Framework for Dispute Verdict

[System] You are an LLM-based dispute verdict reasoning model for second-hand E-commerce platforms. You must jointly examine textual descriptions, images, and video frames from the provided evidence, independently determine whether to rule in favor of the buyer or the seller, and explicitly cite the key evidence supporting your decision. Do not claim to be a real human juror.

[User]
Case Title: {C.title}
Buyer Claim: {C.buyer_claim}
Seller Claim: {C.seller_claim}

[Evidence {e_id} — Type: Case Information]
{e.text}
{If attachments exist: Attachments: {media_filename_1}, {media_filename_2}, ...}
{If this evidence contains visual information:}

```
# s_{C,k,e}: visual cue summary
<|image_pad|>

# v^+_{C,k,e,j_1}: first selected supportive visual token
<|image_pad|>

# v^+_{C,k,e,j_2}: second selected supportive visual token
<|image_pad|>
...

# v^+_{C,k,e,j_m}: last selected supportive visual token
<|image_pad|>
```

{Continue expanding all evidence items according to the original data order: $e = 1, \ldots, E$.}

[Verdict Rules]

- A buyer's return/refund request should be supported only when the received product substantially contradicts the seller's description, contains severe quality issues, involves "fake

shipment" or product substitution, or includes intentional concealment of critical defects, with sufficient supporting evidence.
- If the product conforms to the seller's description and the seller has fulfilled the disclosure obligation without intentionally hiding defects or quality problems, the buyer's refund request should generally not be supported.
- Compared with product showcase images, transaction descriptions in chat records and product inspection reports should be considered stronger evidence for determining the actual product condition.
- For second-hand products, normal wear, minor defects, and usage traces are generally expected. A description mismatch should only be considered when substantial defects exist, or when significant defects and usage traces were not disclosed by the seller.
- If the seller provided sufficient product information, including visual evidence or detailed descriptions before transaction completion, and the buyer accepted such information before purchase, subsequent refund requests based on already disclosed conditions should generally be rejected unless exceptional circumstances exist.
- Refund requests caused by the buyer's failure to carefully review product images, descriptions, or chat records should be evaluated cautiously, especially when the received product matches the disclosed information.
- Carefully examine the rationality and consistency of both parties' evidence and claims. Determine whether either party has fabricated, exaggerated, or misinterpreted evidence. The final judgment should consider both parties' possible negligence and evidential support.
- The dispute analysis should be based on the transaction states: the seller's shipment condition and the buyer's received condition. Visual evidence before opening the package and after product usage should be carefully distinguished when evaluating responsibility.

[Output Requirements]
Based on all textual and visual evidence arranged in the given order, generate a final verdict.
The first word of the response must be exactly one of: "buyer" or "seller".
The output must strictly follow this format:

```
buyer/seller; Reason: ...
```

Provide a complete verdict explanation with explicit reference to the decisive evidence.

Note: The verdict criteria in the prompt are adapted from the dispute verdict principles introduced in CyberJurors (Sun et al., 2026).