ROUTE, FUSE, AND VERIFY: COUNTERFACTUAL EVIDENCE-GUIDED EXPERT ROUTING FOR WHOLE-BODY HUMAN RECOGNITION

Anonymous authors
Paper under double-blind review

ABSTRACT

Multimodal human recognition can exploit complementary face, gait, and body-appearance experts, but existing methods typically adapt only one part of the recognition pipeline, such as model weighting, expert selection, or score calibration. Because these stages rely on separate signals, the system lacks a candidate-specific representation indicating whether the selected models provide sufficient identity evidence and whether the query may come from outside the gallery. As a result, they lack a direct way to determine whether the current evidence is sufficient for an identity decision or whether an additional expert is likely to provide useful identity information. We propose a query-adaptive framework that connects model routing, score fusion, and open-set recognition through calibrated identity evidence. For each selected model subset, the Calibrated Identity Evidence Model (CIEM) estimates candidate genuineness and query unknownness, using their difference as a verification certificate. The certificate guides a multimodal policy optimized with GRPO to select an anchor expert, invoke complementary experts, or stop inference. Its routing reward combines retrieval utility, verification and open-set safety, decision correctness, and the cost of additional expert calls. Given the selected route, Anchor-Preserving Temporal-Consensus (APTC) uses the anchor's complete score vector as the global base and adds high-confidence temporal evidence from complementary experts. Certificate-Reanchored Identity Scoring (CRIS) then projects CIEM certificate margins onto the APTC identity ranking through an order-preserving isotonic transformation. Experiments on CCVID, MEVID, and LTCC show that the resulting pipeline improves retrieval, verification, and open-set identification over fixed fusion, quality-guided fusion, and metric-reward routing baselines.

1 INTRODUCTION

Human biometric recognition integrates complementary cues from facial appearance (Deng et al., 2019; Kim et al., 2022a), gait (Ye et al., 2024; Zhang et al., 2019), and body appearance (Gu et al., 2022; Liu et al., 2024b) to identify individuals under challenging conditions, such as occlusion, low resolution, viewpoint variation, and clothing changes. Different recognition models exhibit complementary strengths and failure modes. Therefore, multimodal human recognition requires not only discriminative feature representations, but also the ability to select models for each query, fuse heterogeneous scores, and accept enrolled identities or reject unknown identities in open-set scenarios.

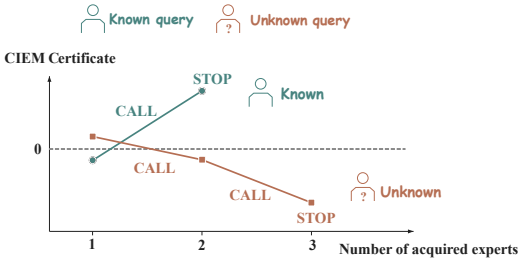

Figure 1: Conceptual illustration of CIEM-guided evidence acquisition. CIEM updates a subset-calibrated certificate margin after each expert call, enabling the policy to acquire additional evidence or stop for identity acceptance or unknown rejection.

Existing multimodal recognition systems mainly rely on either predefined score-fusion

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.

rules or learned query-adaptive fusion and routing strategies. Rule-based methods (Jain et al., 2005; Snelick et al., 2003; Su et al., 2026b) combine model scores using predefined operations. Although they are straightforward and easy to implement, their fixed rules cannot adapt effectively to query-dependent variations in model reliability. Learning-based methods (Park et al., 2021; Teng et al., 2022; Zhu et al., 2025; Ji et al., 2026; Zhu et al., 2026) instead infer model contributions from training data or input quality. With the development of multimodal large language models (Chen et al., 2025; Guo et al., 2026; Yang et al., 2025a; Zheng et al., 2026; Su et al., 2026a), this adaptive paradigm has expanded from learning fusion weights to query-level model selection and sequential expert invocation.

Despite this progress, existing methods make routing and fusion decisions using different criteria, such as input quality, prediction confidence, model agreement, and intermediate recognition results. Agentic approaches can already invoke experts sequentially and decide whether to continue or terminate inference. However, these signals are generally not formulated as an explicit evidence margin that is simultaneously conditioned on a candidate identity and the currently selected model subset, calibrated for open-set acceptance, and reusable for final identity-score calibration. Consequently, existing systems lack a common evidence quantity for evaluating the current model subset, guiding subsequent expert acquisition, and calibrating the final open-set decision.

To address this limitation, we propose the Calibrated Identity Evidence Model (CIEM). Given a query, a candidate identity, and the currently selected model subset, CIEM separately estimates genuine-match support for the candidate and the likelihood that the query belongs to an unenrolled identity. Their difference is compared with a threshold calibrated specifically for the current model subset, producing a certificate margin that indicates whether the current evidence supports accepting the candidate. As conceptually illustrated in Fig. 1, the margin is updated after each expert call: sufficient positive evidence can support identity acceptance, whereas insufficient evidence can trigger another expert call or lead to unknown-identity rejection. CIEM therefore provides a common evidence signal for sequential routing and final open-set calibration.

CIEM evaluates the evidence provided by an existing model subset, but it does not determine how the route should be constructed. We therefore optimize a multi-turn model-routing policy with GRPO. Based on the query content and CIEM evidence accumulated across turns, the policy selects an anchor model, invokes additional models, or terminates inference. The policy reward is aligned with the final APTC fusion outcome. The policy is trained with rewards at two levels. At the route level, the reward evaluates query-level retrieval quality, template- and identity-level ranking margins, CIEM verification evidence, and the cost of additional model calls, and penalizes routes that reduce verification or open-set performance relative to the corresponding ACT route. At the answer level, the policy receives a binary reward when the final identity string exactly matches the ground-truth label. The resulting policy learns an anchor-aware route that balances recognition reliability and inference efficiency, rather than selecting a model combination with strong average performance.

At the final scoring stage, we introduce Anchor-Preserving Temporal-Consensus (APTC). APTC uses the complete sequence-level score vector of the anchor model as the global base score and applies temporal evidence from non-anchor models only to their high-confidence candidates. We then introduce Certificate-Reanchored Identity Scoring (CRIS), which maps the CIEM certificate margins onto the identity ordering established by APTC. CRIS recalibrates the open-set identity scores without changing the identity ranking produced by APTC. Together, CIEM, GRPO, APTC, and CRIS form a unified identity-decision pipeline for evidence evaluation, evidence acquisition, score fusion, and open-set calibration.

Our main contributions are as follows:

- We propose CIEM, which jointly models candidate-identity genuineness and query unknownness for the current model subset, producing a candidate-specific, subset-specific, and open-set-calibrated identity certificate;

- We propose a multi-turn model-routing method jointly supervised by CIEM evidence and APTC outcomes. The method optimizes query-level recognition quality while explicitly constraining false acceptance of unknown identities, false rejection of enrolled identities, and additional model-invocation costs;

- We propose a scoring framework consisting of APTC and CRIS. APTC incorporates complementary evidence from non-anchor models, while CRIS calibrates open-set identity scores without changing the identity ordering produced by APTC.

2 Related Works

2.1 Whole-Body Biometric Recognition

Biometric recognition covers multiple modalities, including face recognition (Deng et al., 2019; Kim et al., 2022a; 2025b;c; You et al., 2025), gait recognition (Ye et al., 2024; Zhang et al., 2019; Habib et al., 2025; Huang et al., 2025; Jin et al., 2025; Ye et al., 2026; Yang et al., 2025b), and person re-identification (Gu et al., 2022; Liu et al., 2024b; Cui et al., 2025; Khalid et al., 2025; Li et al., 2025; Su et al., 2025; Xu et al., 2025; Zhang et al., 2025). When facial information is unclear or captured from an oblique viewpoint (Deng et al., 2020; Kim et al., 2022b), gait models can provide complementary evidence by focusing on body dynamics. Accordingly, most existing studies have focused on designing more effective fusion strategies or using large-model agents for adaptive model invocation to improve overall recognition performance. However, prior methods typically address only one stage of the biometric recognition pipeline, such as input quality estimation or model invocation order, and do not explicitly characterize how much reliable evidence the current query and model subset provide for a candidate identity. We use CIEM to jointly model candidate-identity genuineness and query unknownness, providing a unified evidence representation for subsequent model invocation and open-set identity decisions.

2.2 Score Fusion

Unlike feature-level fusion (Arevalo et al., 2017; Kim et al., 2025a; Wang et al., 2025b;a), score fusion integrates similarity scores from different modalities to improve the final recognition decision. Existing score-fusion methods can be broadly divided into rule-based and learning-based approaches. Rule-based methods typically apply predefined operations, such as Z-score normalization, min-max normalization, or likelihood-ratio-based fusion (Jain et al., 2005; 2006; He et al., 2010). Learning-based methods adaptively optimize the fusion process during training. For example, QME (Zhu et al., 2025) performs quality-guided score fusion by adjusting the contributions of different modalities, while FusionAgent (Zhu et al., 2026) uses the complete score vector of an anchor model and top-k corrections from the selected models. Building on these approaches, we use APTC to combine the global sequence-level scores of the anchor with temporal evidence from non-anchor models. CRIS then calibrates the open-set identity scores while preserving the identity ordering produced by APTC.

2.3 Large-Model Reinforcement Learning

Recent studies have used reinforcement learning to train language and vision-language models to interact with external tools. In these methods, different models, retrievers, or execution modules are treated as tools, and the policy determines the invocation order, number of calls, and termination point based on the input and intermediate results. Group Relative Policy Optimization (GRPO) (Shao et al., 2024) provides a convenient framework for optimizing such policies with group-relative rewards without requiring a separate value model. ToolRL (Qian et al., 2026) shows that reinforcement learning can induce emergent capabilities even when the policy is initialized from a base model without imitation traces, including self-correction of erroneous code, adaptive tool-use frequency, and composition of multiple tools for complex subtasks. More recent studies, such as DeepEyes (Zheng et al., 2026) and Pixel-Reasoner (Su et al., 2026a), further use adaptive tool invocation to improve the performance of vision-language tasks. In our method, CIEM connects the evidence from the selected recognition models to identity decisions, while the APTC-aligned GRPO reward coordinates model routing, final score fusion, and open-set identity recognition.

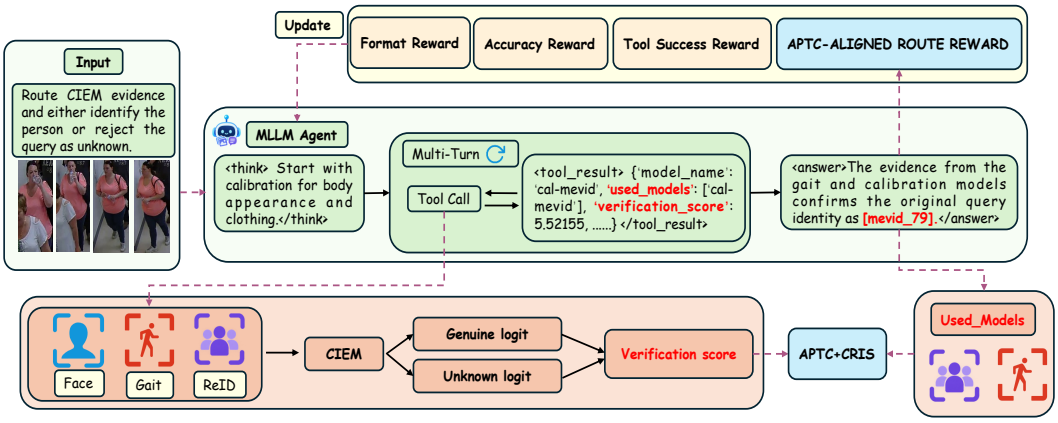

Figure 2: Overview of the proposed CIEM-guided multimodal recognition framework. Given a query, the MLLM agent performs multi-turn calls to frozen face, gait, and ReID experts. CIEM converts the accumulated expert evidence into genuine and unknown logits and a verification score, which guides subsequent model calls or termination. The selected models are then fused by APTC and calibrated by CRIS, while format, tool-success, accuracy and APTC-aligned routing rewards optimize the policy.

3 METHOD

3.1 PROBLEM FORMULATION AND FRAMEWORK OVERVIEW

In biometric recognition, a query (or probe) is an image or video sequence to be compared with an enrolled gallery. Let $q \in \mathcal{Q}$ denote a query,

$$\mathcal{G} = \{g_1, g_2, \ldots, g_{N_g}\}$$

the gallery containing N_g templates, and $\mathcal{Y} = \{y_1, \ldots, y_{N_y}\}$ the set of gallery identities. The system contains a predefined and frozen set of recognition models

$$\mathcal{M} = \{m_1, m_2, \ldots, m_M\}.$$

For a model $m \in \mathcal{M}$, let $u_{m,q}$ denote the sequence-level feature extracted from query q, and let $v_{m,j}$ denote the feature of gallery template g_j. Each model uses its native similarity function sim_m to produce a query-to-gallery score vector

$$\mathbf{s}_{m,q} = [s_{m,q,1}, \ldots, s_{m,q,N_g}], \qquad s_{m,q,j} = \mathrm{sim}_m(u_{m,q}, v_{m,j}).$$

The query q always refers to the same input sequence; feature aggregation is performed inside each recognition model. For cosine-based models, sim_m is a normalized inner product. For distance-based models, distance is converted into a similarity that increases monotonically with match likelihood. Consequently, all model outputs have the same interpretation: a larger score indicates that the query and gallery template are more likely to share an identity.

Following FusionAgent, recognition consists of query-level model selection followed by query-level score fusion. Our policy produces an ordered route rather than an unordered model subset:

$$\rho_q = (m_{(1)}, m_{(2)}, \ldots, m_{(L)}), \qquad \mathcal{M}_q = \{m_{(1)}, \ldots, m_{(L)}\} \subseteq \mathcal{M}.$$

The first invoked model $a_q = m_{(1)}$ is the anchor, and L is the number of model calls. Given the selected models, a fusion function maps their score vectors to the final template score:

$$\mathbf{s}'_q = F(\{\mathbf{s}_{m,q} : m \in \mathcal{M}_q\}, a_q) \in \mathbb{R}^{N_g}.$$

With three models, the anchor-aware route space contains 12 routes: three single-model routes, six ordered two-model routes, and three full-model routes with different anchors.

The proposed framework has three components. First, CIWM produces an open-set certificate for each candidate identity from the evidence of the currently selected model subset. Second, a multi-modal policy uses the returned certificate to call another model or stop; the policy is optimized with GRPO using a reward aligned with the final APTC result. Third, APTC fuses the template scores of the selected models, after which CRIS uses the CIWM certificate to calibrate the open-set identity scores.

3.2 CALIBRATED IDENTITY EVIDENCE MODEL

Identity explanation energy. For expert m, query embedding x_m, and candidate support set $S_{m,c} = \{s_{m,c,k}\}_{k=1}^{K_c}$, let $p^0_{m,c}$ be the normalized pooled support prototype. The raw residual is

$$E^0_m(q,c) = 1 - \langle \mathrm{norm}(x_m), p^0_{m,c} \rangle.$$

A learned encoder and camera/clothing conditioner produce a conditioned prototype $\widehat{p}_{m,c}$ and residual $E^{\mathrm{cond}}_m(q,c) = 1 - \langle z^q_m, \widehat{p}_{m,c} \rangle$. With support dispersion $D_m(c)$, the bounded correction is

$$E_m(q,c) = (1-\alpha_m)E^0_m(q,c) + \alpha_m\big(E^{\mathrm{cond}}_m(q,c) + 0.05D_m(c)\big), \qquad 0 < \alpha_m \le 0.25.$$

Lower energy indicates stronger support. A pairwise softplus loss ranks the genuine candidate below sampled negatives while regularizing α_m.

Genuine–unknown certificate. For route R, CIEM stacks model–identity energies into a matrix; unselected or unavailable entries are median-imputed while a binary availability channel is retained. Candidate features summarize per-model energy, competitor gap, relative rank, and availability, whereas query features summarize the candidate-energy distribution. Two lightweight heads output

$$\ell_g(q,c;R) = h_g(x_c), \qquad \ell_u(q;R) = h_u(x_q), \qquad V(q,c;R) = \ell_g(q,c;R) - \ell_u(q;R).$$

The certifier is trained on genuine candidates, hard impostors, and identity-disjoint pseudo-unknown queries using weighted classification, unknown-rejection, and genuine–impostor margin terms. For each route, an OOF threshold τ_R is calibrated at 1% FPIR. CIEM predicts

$$\widehat{y}_R(q) = \arg\max_c V(q,c;R), \qquad d_R(q) = \max_c V(q,c;R) - \tau_R,$$

and accepts $\widehat{y}_R(q)$ iff $d_R(q) \ge 0$. The policy receives the current route, logits, margin, identity, acceptance state, and unused experts.

3.3 ORDERED ROUTING WITH GRPO

Interaction. At each turn, the policy calls one unused expert or stops. The first legal call fixes the anchor; repeated or unavailable calls and empty routes are invalid. Training and inference use this same state machine, with at most M calls.

3.3.1 APTC-ALIGNED ROUTE REWARD

All feasible anchor-aware routes are evaluated using train-only out-of-fold (OOF) supervision. For query q and route r, we define the query-level retrieval utility as

$$U(q,r) = \frac{1}{2}\,\mathrm{AP}_{q,r} + \frac{1}{2}\,\mathrm{RR}_{q,r}\,. \tag{1}$$

Let $\mathcal{R}(q)$ denote the feasible route set and $U_q = \{U(q,r') : r' \in \mathcal{R}(q)\}$. We robustly normalize the utility within each query:

$$Z_q(U) = \mathrm{clip}\left(\frac{U(q,r) - \mathrm{median}(U_q)}{\max\left(\mathrm{MAD}(U_q), \mathrm{std}(U_q), 0.05\right)}, -1, 1\right). \tag{2}$$

Because APTC modifies the template branch, we use its route-level OOF TAR as a weak ACT-relative safety prior:

$$\Delta_{\mathrm{TAR}}(r) = \mathrm{TAR}_{\mathrm{APTC}}(r) - \mathrm{TAR}_{\mathrm{ACT}}(r). \tag{3}$$

The corresponding TAR violation indicator is

$$V_{\mathrm{TAR}}(r) = \mathbf{1}\left[\Delta_{\mathrm{TAR}}(r) < -0.002\right]. \tag{4}$$

This is a soft template-safety regularizer rather than a hard route filter. We do not include a pre-CRIS FNIR term because APTC recomputes identity scores using the same pooled ACT identity branch.

Therefore, under an identical route, APTC and ACT have the same pre-CRIS identity matrix and FNIR, so their FNIR difference provides no route-discriminative reward signal.

CIEM supplies an auxiliary decision indicator

$$D(q,r)=\begin{cases}1, & q \text{ is known, accepted, and correctly identified,}\\ 1, & q \text{ is pseudo-unknown and correctly rejected,}\\ 0, & \text{otherwise.}\end{cases} \tag{5}$$

For a legal route, the scalar route reward is

$$\begin{aligned}R_{\text{route}}(q,r)=\text{clip}\Big(&Z_q(U)+0.10U(q,r)+0.05D(q,r)\\&-0.05V_{\text{TAR}}(r)-0.05\max(0,|r|-1),-1,1\Big).\end{aligned} \tag{6}$$

Invalid, unavailable, or empty routes receive $R_{\text{route}}=-1$. The complete scalar reward is

$$R(q,r)=0.10R_{\text{format}}+0.10R_{\text{tool}}+R_{\text{route}}+0.25R_{\text{answer}}. \tag{7}$$

Here, R_{answer} equals one only when the final predicted identity exactly matches the ground-truth identity. The same reward coefficients are used for CCVID, MEVID, and LTCC. In the prefix-marginal variant, the same scalar route value is evaluated on successive route prefixes and differenced to provide turn-level credit; the scalar route objective itself remains unchanged.

3.3.2 CRIS and Final FNIR Evaluation

APTC uses temporal evidence only in the template branch. Its identity branch is recomputed from the original pooled expert scores:

$$S_{\text{id}}^{\text{APTC}}(q,y;r)=S_{\text{id}}^{\text{ACT}}(q,y;r). \tag{8}$$

Consequently, pre-CRIS FNIR is structurally unchanged and is not included in the route reward.

For the selected route R, CIEM provides route-calibrated certificate margins

$$M_R(q,y_c)=V(q,c;R)-\tau_R, \tag{9}$$

where τ_R is calibrated on identity-disjoint OOF pseudo-unknown queries at the target FPIR.

Let π_q be the descending APTC identity order and define $c_i=M_R(q,y_{\pi_q(i)})$. CRIS performs the isotonic projection

$$\mathbf{u}^*=\arg\min_{u_1\geq\cdots\geq u_{N_y}}\sum_{i=1}^{N_y}(u_i-c_i)^2,\qquad S_{\text{CRIS}}(q,y_{\pi_q(i)})=u_i^*. \tag{10}$$

CRIS changes the calibrated spacing between identity scores while preserving the APTC identity order. Therefore, template retrieval and verification are evaluated using S_{APTC}, whereas open-set identity search and FNIR are evaluated using S_{CRIS}. The resulting FNIR improvement is measured empirically; rank preservation alone does not imply a universal monotonic FNIR guarantee.

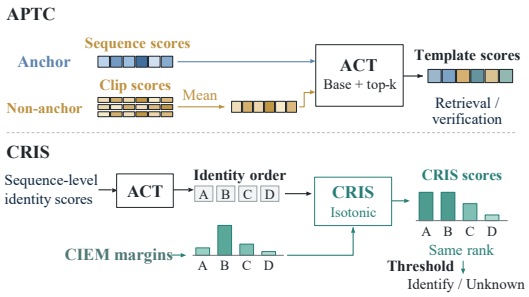

Figure 3: Illustration of APTC and CRIS.

4 Experiments

4.1 Experimental Setup

Datasets and Evaluation Metrics. We evaluate our method on three whole-body biometric

benchmarks: CCVID (Gu et al., 2022), MEVID Davila et al. (2023), and LTCC Qian et al. (2020). These datasets provide complementary evaluation settings involving video- and image-based recognition, viewpoint variation, low-quality observations, and clothing changes. Together, they allow us to assess closed-set retrieval, verification, open-set identification, and the robustness of query-dependent model routing. Following the evaluation protocol of FusionAgent (Zhu et al., 2026), we report Rank-1 accuracy and mAP for retrieval, TAR@1% FAR for verification, and FNIR@1% FPIR for open-set identification. For datasets with clothing-aware evaluation protocols, we additionally follow the corresponding standard SC and CC evaluation settings when these metrics are reported.

Implementation Details. We conduct the training and evaluation processes on four NVIDIA RTX 3090 GPUs. We use the same frozen biometric model pools as the corresponding benchmark protocols. Each expert produces query and gallery representations in its native feature space. For the CIEM evidence model, we construct identity-conditioned explanation energies from query features and identity support features. For APTC, the anchor uses the sequence-level score vector, while non-anchor experts provide complementary scores obtained by averaging similarities between orbit tokens and gallery templates. The CIEM certifier and all biometric experts remain frozen during GRPO training. The routing policy is initialized from Qwen2.5-VL-3B-Instruct (Bai et al., 1) and fine-tuned with GRPO. We use LoRA (Hu et al., 2021) with rank $r = 64$ and scaling factor $\alpha = 128$. The learning rate is set to 5×10^{-6}, and training runs for 200 optimization steps with gradient checkpointing enabled. The policy samples six trajectories per query. The maximum number of interaction turns is four, and each turn allows at most one expert call or a STOP action. The APTC top-k parameter is set to $k = 10$ for CCVID and $k = 40$ for MEVID and LTCC. The final routing reward is computed from the APTC route utility, CIEM decision correctness, route-level TAR/FNIR safety constraints, and the cost of additional expert calls. All comparisons use the same gallery construction, score protocol, and open-set evaluation procedure as the corresponding baseline.

4.2 Comparisons with Prior Work

In-domain Evaluation. Table 1 compares multiple individual models, multimodal model combinations, and score-fusion methods. The performance of a single model can vary substantially across evaluation conditions. For example, AdaFace achieves 94.0% Rank-1 accuracy on CCVID, but its Rank-1 accuracy drops to 18.5% on LTCC. Statistical fusion baselines, such as Z-score normalization and FarSight, exploit complementary information from different modalities to some extent, but their overall performance remains limited. Learning-based methods, including QME and FusionAgent, further improve recognition performance, but still exhibit relatively high FNIR in open-set identification. Our method improves Rank-1, mAP, and TAR while substantially reducing FNIR. We attribute these gains to CIEM, which explicitly models candidate-identity genuineness and query unknownness, together with APTC, which integrates complementary evidence from the selected models to improve performance across all four metrics.

Cross-domain Evaluation. Real-world systems may encounter acquisition conditions and data distributions absent from the training data, making cross-domain performance important for evaluating model generalization and reliability. The CIEM identity explanation energy and APTC route supervision depend on identity support samples from the target domain, making strict zero-shot evaluation infeasible. We therefore train the model on CCVID and conduct 10-shot evaluations on MEVID and LTCC. As shown in Table 2, our method demonstrates strong generalization to both target datasets. These results suggest that the model does not merely memorize fixed routing patterns from the source domain. Instead, candidate-identity evidence, query unknownness, and model availability captured by CIEM provide stable signals for routing decisions across different environments.

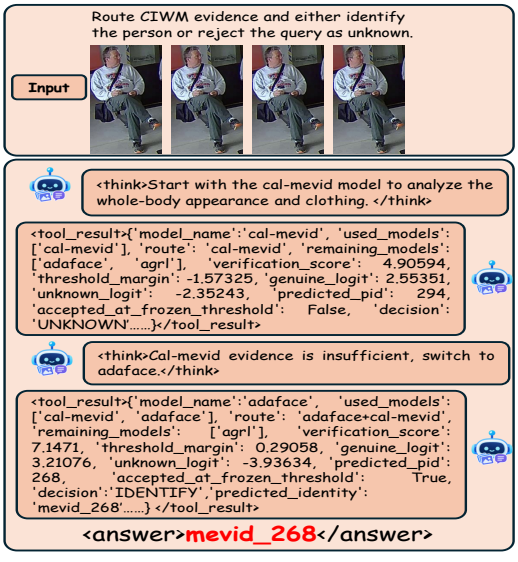

Figure 4: Examples of Qualitative results.

Table 1: Comparison with the state-of-the-art methods on the CCVID, MEVID and LTCC dataset. Best and second best performance.

Method	CCVID				MEVID				LTCC			
	Rank1↑	mAP↑	TAR↑	FNIR↓	Rank1↑	mAP↑	TAR↑	FNIR↓	Rank1↑	mAP↑	TAR↑	FNIR↓
Adaface (Kim et al., 2022a)	94.0	87.9	75.7	13.0±3.5	25.0	8.1	5.4	98.8±1.2	18.5	5.9	2.4	99.8±0.2
CAL (Gu et al., 2022)	84.1	74.7	60.3	52.8+13.3	52.5	27.1	34.7	67.8+7.3	74.4	40.6	36.7	59.7±7.3
Biggait (Ye et al., 2024)	78.7	61.0	49.7	71.1±6.1	-	-	-	-	-	-	-	-
AGRL (Wu et al., 2020)	-	-	-	-	51.9	25.5	30.7	69.4±8.9	-	-	-	-
AIM (Yang et al., 2023)	-	-	-	-	-	-	-	-	74.8	40.9	37.0	66.2±7.5
SapiensID (Kim et al., 2025a)	92.6	77.8	-	-	-	-	-	-	72.0	34.6	-	-
Min-Fusion (Jain et al., 2005)	87.1	79.2	62.4	48.5±8.7	46.8	21.1	28.0	70.4±8.0	38.1	13.5	12.4	81.9±6.0
Max-Fusion (Jain et al., 2005)	89.9	89.3	73.4	23.0±10.1	33.2	14.9	8.3	97.4±1.6	62.5	33.3	16.8	94.8±4.7
Z-score (Snelick et al., 2003)	92.2	90.6	73.9	15.1±1.5	54.1	27.4	30.7	66.5±7.0	73.0	37.5	30.4	68.7±9.2
Min-max (Snelick et al., 2003)	91.8	90.9	73.9	15.4±2.5	52.8	24.7	25.0	71.3±6.1	73.2	38.1	31.9	75.1+9.2
Weighted-sum (Park et al., 2021)	91.7	90.6	73.6	15.4±1.8	54.1	27.3	30.3	66.3±7.0	73.2	37.8	31.3	72.4±8.6
Asym-AO1 (Herbadji et al., 2020)	92.3	90.0	74.0	15.9±1.7	52.5	22.9	23.6	71.7±5.8	71.2	32.9	19.1	76.3±8.9
BSSF (Teng et al., 2022)	91.8	91.1	73.9	14.1±1.3	53.5	27.4	30.5	65.9±7.2	73.5	39.1	34.2	68.9±8.5
Farsight (Liu et al., 2024a)	92.0	91.2	73.9	13.9±1.1	53.8	25.4	26.6	69.8±6.4	73.2	37.8	31.3	72.4±8.6
QME (Zhu et al., 2025)	**94.1**	90.8	76.2	12.3±1.4	55.7	28.2	32.9	64.6±8.2	73.8	39.6	35.0	64.3±8.0
Fusionagent (Zhu et al., 2026)	93.4	**92.6**	85.9	10.1±1.5	54.7	28.7	34.9	58.6±7.4	75.5	41.0	**37.0**	50.0±8.5
Ours	93.9	92.3	**86.6**	**9.3±1.2**	**57.6**	**30.1**	**35.8**	**52.7±6.0**	**76.5**	**41.3**	36.9	**49.3±8.3**

Table 2: Cross-domian performance on LTCC.

Method	$CCVID \rightarrow LTCC(10-shot)$				$MEVID \rightarrow LTCC(10-shot)$				$CCVID \rightarrow MEVID(10-shot)$			
	Rank1↑	mAP↑	TAR↑	FNIR↓	Rank1↑	mAP↑	TAR↑	FNIR↓	Rank1↑	mAP↑	TAR↑	FNIR↓
Fusionagent	73.6	39.8	34.8	53.5±8.5	-	-	-	-	-	-	-	-
Ours	**74.8**	**40.9**	**36.9**	**51.1±6.2**	75.5	41.0	34.5	49.8±7.6	53.2	27.3	35.4	54.4±5.5

Qualitative Results. Fig.4 illustrates a sequential recognition process for a query from MEVID. The policy first invokes the cal-mevid model. Although this model produces a verification score of 4.90594, its threshold margin is -1.57325, indicating that the current evidence is insufficient for identity acceptance. The system temporarily returns UNKNOWN and acquires additional evidence. The policy then invokes AdaFace. For the combined route, the verification score increases to 7.14717, and the threshold margin becomes 0.29058, exceeding the acceptance threshold. The query is consequently identified as $mevid_268$. This example demonstrates that CIEM can determine whether the evidence provided by the current model subset is sufficient, while GRPO selects a complementary expert when additional evidence is needed. The final APTC/CRIS pipeline then uses the combined evidence to make identity and open-set decision.

4.3 Ablation Study

APTC and CRIS. Under the same routing strategy, model pool, and evaluation protocol, we compare ACT, APTC, and APTC+CRIS. The results show that APTC consistently improves retrieval and verification by preserving the anchor's global score structure while incorporating complementary temporal evidence from non-anchor models. Adding CRIS further uses the CIEM certificates to recalibrate identity scores while preserving their ranking, thereby improving open-set identity acceptance and unknown-identity rejection. APTC+CRIS achieves the best overall performance, demonstrating that APTC and CRIS play complementary roles: APTC integrates complementary evidence, whereas CRIS improves open-set calibration without altering the identity ordering.

Route-reward component ablation. Table 4 evaluates the two auxiliary components of the route reward on CCVID. "w/o db" removes the CIEM decision bonus, while "w/o ta" removes the route regularization terms, including the TAR safety penalty and the additional-expert cost. Removing the decision bonus reduces Rank-1 accuracy from 93.9% to 92.3%, mAP from 92.3% to 90.7%, and TAR from 86.6% to 48.4%, while increasing FNIR from 9.3% to 10.4%. Removing the route regularization causes a larger degradation, reducing Rank-1 accuracy to 92.9% and mAP to 91.4%, and increasing FNIR to 10.5%. These results show that the CIEM decision bonus provides useful certificate-aware supervision beyond retrieval utility, while route regularization is important for maintaining reliable verification and open-set recognition. The full route reward achieves the best performance across all reported metrics.

5 Conclusion

This paper proposes a query-adaptive framework for multimodal human recognition. CIEM jointly models candidate identity authenticity and query unknownness to provide open-set identity evidence for the current model subset; GRPO uses this evidence to select an anchor model, invoke complementary experts, or terminate inference. APTC fuses global and temporal information from the selected models, while CRIS calibrates open-set identity scores while preserving identity rank-

Table 3: Ablation results of APTC and CRIS on CCVID.

Method	Rank1↑	mAP↑	TAR↑	FNIR↓
ACT	92.9	92.2	85.7	11.8 ± 1.47
APTC	93.9	92.3	86.6	11.8 ± 1.47
APTC+CRIS	**93.9**	**92.3**	**86.6**	**9.3±1.2**

Table 4: Ablation of route-reward components on CCVID.

Method	Rank1↑	mAP↑	TAR↑	FNIR↓
w/o db	92.3	90.7	84.4	10.4 ± 1.3
w/o ta	92.9	91.4	84.9	10.5 ± 1.5
R_full	**93.9**	**92.3**	**86.6**	**9.3±1.2**

ing. Experiments on CCVID, MEVID, and LTCC demonstrate that the proposed method improves recognition and verification performance while enhancing open-set identity search. These results suggest that connecting identity evidence, model routing, score fusion, and open-set calibration can facilitate more reliable multimodal human recognition systems.

6 AI USE STATEMENT

In this work, we used generative AI tools for language polishing, figure refinement, and generating parts of the code. All AI-assisted text and figures were reviewed by the authors, and the AI-generated code was reviewed by one author for correctness. We take responsibility for the final content of this work, including text, claims, figures, and code produced with the aid of generative AI.

REFERENCES

John Arevalo, Thamar Solorio, Manuel Montes-y Gómez, and Fabio A González. Gated multimodal units for information fusion. *arXiv preprint arXiv:1702.01992*, 2017.

Shuai Bai, Keqin Chen, Xuejing Liu, Jialin Wang, Wenbin Ge, Sibo Song, Kai Dang, Peng Wang, Shijie Wang, Jun Tang, et al. others. 2025. qwen2. 5-vl technical report. *arXiv preprint arXiv:2502.13923*, 4(5):1, 1.

Boyu Chen, Zhengrong Yue, Siran Chen, Zikang Wang, Yang Liu, Peng Li, and Yali Wang. Lvagent: Long video understanding by multi-round dynamical collaboration of mllm agents. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 20237–20246. IEEE, 2025.

Zhenyu Cui, Jiahuan Zhou, and Yuxin Peng. Dkc: Differentiated knowledge consolidation for cloth-hybrid lifelong person re-identification. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3573–3582. IEEE, 2025.

Daniel Davila, Dawei Du, Bryon Lewis, Christopher Funk, Joseph Van Pelt, Roderic Collins, Kellie Corona, Matt Brown, Scott McCloskey, Anthony Hoogs, et al. Mevid: Multi-view extended videos with identities for video person re-identification. In *2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pp. 1634–1643. IEEE, 2023.

Jiankang Deng, Jia Guo, Niannan Xue, and Stefanos Zafeiriou. Arcface: Additive angular margin loss for deep face recognition. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 4690–4699, 2019.

Jiankang Deng, Jia Guo, Tongliang Liu, Mingming Gong, and Stefanos Zafeiriou. Sub-center arcface: Boosting face recognition by large-scale noisy web faces. In *European Conference on Computer Vision*, pp. 741–757. Springer, 2020.

Xinqian Gu, Hong Chang, Bingpeng Ma, Shutao Bai, Shiguang Shan, and Xilin Chen. Clothes-changing person re-identification with rgb modality only. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1050–1059. IEEE, 2022.

Xiao Guo, Yue Zhang, Mohit Bansal, and Xiaoming Liu. Deepfake-agent: Aggregating semantic forgery clues for generalizable detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4898–4908, 2026.

Gavriel Habib, Noa Barzilay, Or Shimshi, Rami Ben-Ari, and Nir Darshan. Cargait: Cross-attention based re-ranking for gait recognition. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 11884–11894. IEEE, 2025.

Mingxing He, Shi-Jinn Horng, Pingzhi Fan, Ray-Shine Run, Rong-Jian Chen, Jui-Lin Lai, Muhammad Khurram Khan, and Kevin Octavius Sentosa. Performance evaluation of score level fusion in multimodal biometric systems. *Pattern Recognition*, 43(5):1789–1800, 2010.

Abderrahmane Herbadji, Zahid Akhtar, Kamran Siddique, Noubeil Guermat, Lahcene Ziet, Mohamed Cheniti, and Khan Muhammad. Combining multiple biometric traits using asymmetric aggregation operators for improved person recognition. *Symmetry*, 12(3):444, 2020.

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. *arXiv preprint arXiv:2106.09685*, 2021.

Panjian Huang, Saihui Hou, Junzhou Huang, and Yongzhen Huang. Learning a unified template for gait recognition. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 12459–12469. IEEE, 2025.

Anil Jain, Karthik Nandakumar, and Arun Ross. Score normalization in multimodal biometric systems. *Pattern recognition*, 38(12):2270–2285, 2005.

Anil K Jain, Ruud Bolle, and Sharath Pankanti. *Biometrics: personal identification in networked society*. Springer Science & Business Media, 2006.

Yuyang Ji, Yixuan Shen, Kien Nguyen, Lifeng Zhou, and Feng Liu. Idselect: A rl-based cost-aware selection agent for video-based multi-modal person recognition. *arXiv preprint arXiv:2602.18990*, 2026.

Dongyang Jin, Chao Fan, Jingzhe Ma, Jingkai Zhou, Weihua Chen, and Shiqi Yu. On denoising walking videos for gait recognition. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 12347–12357. IEEE, 2025.

Wajahat Khalid, Bin Liu, Xulin Li, Muhammad Waqas, and Muhammad Sher Afgan. Bridging the sky and ground: Towards view-invariant feature learning for aerial-ground person reidentification. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 9749–9758. IEEE, 2025.

Minchul Kim, Anil K Jain, and Xiaoming Liu. Adaface: Quality adaptive margin for face recognition. In *2022 IEEE/CVF conference on computer vision and pattern recognition (CVPR)*, pp. 18729–18738. IEEE, 2022a.

Minchul Kim, Feng Liu, Anil K Jain, and Xiaoming Liu. Cluster and aggregate: Face recognition with large probe set. *Advances in Neural Information Processing Systems*, 35:36054–36066, 2022b.

Minchul Kim, Dingqiang Ye, Yiyang Su, Feng Liu, and Xiaoming Liu. Sapiensid: Foundation for human recognition. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 13937–13947. IEEE, 2025a.

Minsoo Kim, Min-Cheol Sagong, Gi Pyo Nam, Junghyun Cho, and Ig-Jae Kim. Vigface: Virtual identity generation for privacy-free face recognition dataset. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 10043–10053. IEEE, 2025b.

Sunpill Kim, Seunghun Paik, Chanwoo Hwang, Dongsoo Kim, Junbum Shin, and Jae Hong Seo. Idface: Face template protection for efficient and secure identification. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 13995–14005. IEEE, 2025c.

Longhua Li, Lei Qi, and Xin Geng. One-shot knowledge transfer for scalable person reidentification. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 668–677. IEEE, 2025.

Feng Liu, Ryan Ashbaugh, Nicholas Chimitt, Najmul Hassan, Ali Hassani, Ajay Jaiswal, Minchul Kim, Zhiyuan Mao, Christopher Perry, Zhiyuan Ren, et al. Farsight: A physics-driven whole-body biometric system at large distance and altitude. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 6227–6236, 2024a.

Feng Liu, Minchul Kim, Zhiyuan Ren, and Xiaoming Liu. Distilling clip with dual guidance for learning discriminative human body shape representation. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 256–266. IEEE, 2024b.

Tae Jin Park, Manoj Kumar, and Shrikanth Narayanan. Multi-scale speaker diarization with neural affinity score fusion. In *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 7173–7177. IEEE, 2021.

Cheng Qian, Emre Can Acikgoz, Qi He, Hongru Wang, Xiusi Chen, Dilek Hakkani-Tur, Gokhan Tur, and Heng Ji. Toolrl: Reward is all tool learning needs. *Advances in Neural Information Processing Systems*, 38:105523–105553, 2026.

Xuelin Qian, Wenxuan Wang, Li Zhang, Fangrui Zhu, Yanwei Fu, Tao Xiang, Yu-Gang Jiang, and Xiangyang Xue. Long-term cloth-changing person re-identification. In *Asian Conference on Computer Vision*, pp. 71–88. Springer, 2020.

Zhihong Shao, Peiyi Wang, Qihao Zhu, Runxin Xu, Junxiao Song, Xiao Bi, Haowei Zhang, Mingchuan Zhang, YK Li, Yang Wu, et al. Deepseekmath: Pushing the limits of mathematical reasoning in open language models. *arXiv preprint arXiv:2402.03300*, 2024.

Robert Snelick, Mike Indovina, James Yen, and Alan Mink. Multimodal biometrics: issues in design and testing. In *Proceedings of the 5th international conference on Multimodal interfaces*, pp. 68–72, 2003.

Alex Su, Haozhe Wang, Weiming Ren, Fangzhen Lin, and Wenhu Chen. Pixel reasoner: Incentivizing pixel space reasoning via curiosity-driven reinforcement learning. *Advances in Neural Information Processing Systems*, 38:8222–8251, 2026a.

Yiyang Su, Yunping Shi, Feng Liu, and Xiaoming Liu. Hamobe: Hierarchical and adaptive mixture of biometric experts for video-based person reid. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 11525–11536. IEEE, 2025.

Yiyang Su, Minchul Kim, Jie Zhu, Christopher Perry, Feng Liu, Anil Jain, and Xiaoming Liu. Localscore: Local density-aware similarity scoring for biometrics. *arXiv preprint arXiv:2602.01012*, 2026b.

Jackson Horlick Teng, Thian Song Ong, Tee Connie, Kalaiarasi Sonai Muthu Anbananthen, and Pa Pa Min. Optimized score level fusion for multi-instance finger vein recognition. *Algorithms*, 15(5):161, 2022.

Langyu Wang, Bingke Zhu, Yingying Chen, and Jinqiao Wang. Link: Adaptive modality interaction for audio-visual video parsing. In *ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1–5. IEEE, 2025a.

Langyu Wang, Bingke Zhu, Yingying Chen, Yiyuan Zhang, Ming Tang, and Jinqiao Wang. Mug: Pseudo labeling augmented audio-visual mamba network for audio-visual video parsing. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 20637–20646. IEEE, 2025b.

Yiming Wu, Omar El Farouk Bourahla, Xi Li, Fei Wu, Qi Tian, and Xue Zhou. Adaptive graph representation learning for video person re-identification. *IEEE Transactions on Image Processing*, 29:8821–8830, 2020.

Kunlun Xu, Fan Zhuo, Jiangmeng Li, Xu Zou, and Jiahuan Zhou. Self-reinforcing prototype evolution with dual-knowledge cooperation for semi-supervised lifelong person re-identification. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 3564–3574. IEEE, 2025.

An Yang, Anfeng Li, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Gao, Chengen Huang, Chenxu Lv, et al. Qwen3 technical report. *arXiv preprint arXiv:2505.09388*, 2025a.

Shaopeng Yang, Jilong Wang, Saihui Hou, Xu Liu, Chunshui Cao, Liang Wang, and Yongzhen Huang. Bridging gait recognition and large language models sequence modeling. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3460–3469. IEEE, 2025b.

Zhengwei Yang, Meng Lin, Xian Zhong, Yu Wu, and Zheng Wang. Good is bad: Causality inspired cloth-debiasing for cloth-changing person re-identification. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1472–1481. IEEE, 2023.

Dingqiang Ye, Chao Fan, Jingzhe Ma, Xiaoming Liu, and Shiqi Yu. Biggait: Learning gait representation you want by large vision models. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 200–210. IEEE, 2024.

Dingqiang Ye, Chao Fan, Zhanbo Huang, Chengwen Luo, Jianqiang Li, Shiqi Yu, and Xiaoming Liu. Biggergait: Unlocking gait recognition with layer-wise representations from large vision models. *Advances in neural information processing systems*, 38:73569–73590, 2026.

Jinghan You, Shanglin Li, Yuanrui Sun, Jiangchuan Wei, Mingyu Guo, Chao Feng, and Jiao Ran. Lvface: Progressive cluster optimization for large vision models in face recognition. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 11840–11849. IEEE, 2025.

Xiao-Wen Zhang, Delong Zhang, Yi-Xing Peng, Zhi Ouyang, Jingke Meng, and Wei-Shi Zheng. Viperson: Flexibly generating virtual identity for person re-identification. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 23374–23384. IEEE, 2025.

Ziyuan Zhang, Luan Tran, Xi Yin, Yousef Atoum, Xiaoming Liu, Jian Wan, and Nanxin Wang. Gait recognition via disentangled representation learning. In *2019 IEEE/CVF conference on computer vision and pattern recognition (CVPR)*, pp. 4705–4714. IEEE, 2019.

Ziwei Zheng, Minghao Yang, Jack Hong, Chenxiao Zhao, Guohai Xu, Le Yang, and Chao Shen. Deepeyes: Incentivizing" thinking with images" via reinforcement learning. In *International Conference on Learning Representations*, volume 2026, pp. 126775–126798, 2026.

Jie Zhu, Yiyang Su, Minchul Kim, Anil Jain, and Xiaoming Liu. A quality-guided mixture of score-fusion experts framework for human recognition. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 13076–13086. IEEE, 2025.

Jie Zhu, Xiao Guo, Yiyang Su, Anil Jain, and Xiaoming Liu. Fusionagent: A multimodal agent with dynamic model selection for human recognition. *arXiv preprint arXiv:2603.26908*, 2026.

A Appendix

A.1 System Prompt for MEVID

This prompt defines the agent as a constrained multimodal biometric analyzer that obtains complementary facial, gait, and body-appearance evidence by calling AdaFace, AGRL, and cal-mevid. The agent follows a CIEM-guided sequential decision process: the first invoked model serves as the anchor, and at each subsequent turn the policy may call one unused model or output STOP, with at most three model calls allowed. After each call, the policy decides whether to acquire additional evidence based on the returned candidate identity, acceptance status, certificate margin, and next-action recommendation. It is not allowed to invent identities or scores. The prompt also specifies a strict output format. For an accepted identity, the agent must copy the complete predicted_identity returned by the tool; for insufficient evidence or an unknown query, it must output unknown. Thus, the prompt organizes model invocation, evidence assessment, stopping decisions, and final identity reporting into an executable state machine.

System prompt

 # Role & Objective
You are an expert-level biometric analysis agent for MEVID. Your primary mission is to achieve the highest possible identification performance by strategically analyzing an input video and selecting the optimal combination of biometric models. Your final answer is a fused identity prediction based only on selected CIWM evidence. There are exactly two outcomes: a supported MEVID identity or unknown.
 # Loop
Use the COT evidence state machine. At each evidence turn emit exactly one CALL_ADAFACE, CALL_AGRL, CALL_CAL_MEVID, or STOP. The first valid CALL selects the APTC anchor and later CALLs add complementary template evidence. Select the anchor from the visible query and returned evidence; do not default every query to one fixed model. An accepted certificate permits an answer but does not force stopping when an unused expert may improve retrieval. After a tool result, call one unused expert or emit a terminal answer. Never emit ROUTE_* or multiple tool calls. Never call a model twice.
 # Strict Output Format
Each turn is either <think>...</think>CALL_MODEL, or a terminal answer. For an identified person, copy the exact predicted_identity string from the latest accepted tool result inside <answer>...</answer>; otherwise output <answer>unknown</answer>.
 # Tag Rules
The evidence turn must summarize only returned CIWM fields: candidate, acceptance, margin/uncertainty, agreement, and next_action. Never invent evidence, identities, or scores.
 # Model Type
adaface handles clear frontal face; agrl handles gait/body motion; cal-mevid handles whole-body appearance and clothing. Legal COT calls are CALL_ADAFACE, CALL_AGRL, CALL_CAL_MEVID, and STOP.
 # Stopping Condition
End with the exact predicted_identity returned by accepted evidence; otherwise end with <answer>unknown</answer>. Never output a schema placeholder or invent an identity.
State-machine example
<think>Start with gait evidence.</think>
CALL_AGRL
<think>The returned evidence supports the accepted identity.</think>
After accepted evidence, copy its exact predicted_identity into the answer tag.
 # Unknown example
<think>Evidence remains insufficient.</think>
STOP
<think>The evidence does not support a database identity.</think>
<answer>unknown</answer>
 # CIWM Evidence State Machine
Available modalities: [adaface, agrl, cal-mevid]. Maximum calls: 3.
adaface: clear frontal face and facial details. agrl: gait sequence and body motion. cal-mevid: whole-body appearance, shape and clothing.
At each turn emit exactly one CALL_<MODEL> action or STOP. Wait for the tool result. After a result, call one unused model or emit a terminal answer. Never emit ROUTE_* or multiple tool calls in one turn; never repeat a model.

Figure 5: System prompt for MEVID.

A.2 More Qualitative Results.

Fig. 6 presents additional CoT examples.The left example illustrates a query for which the initial evidence is insufficient. Although the initial ‘cal-mevid‘ model produces a relatively high verification score, its certificate margin remains below the acceptance threshold calibrated for the current model subset. CIEM therefore does not accept the identity immediately; instead, it identifies the evidence as insufficient and triggers an additional model call. The policy then selects the complementary ‘AGRL‘ model. The combined evidence increases the certificate margin above zero, and the system correctly identifies the query as ‘mevid_211‘. This example demonstrates that our method can determine whether the current identity evidence is sufficient and dynamically decide whether further expert invocation is necessary, thereby avoiding premature acceptance based on inadequate evidence.

The right example shows a query that can be recognized more confidently. ‘AdaFace‘ first provides strong facial identity evidence, with a verification score already exceeding the open-set acceptance threshold. The policy then invokes ‘CAL-CCVID‘ to obtain complementary body-appearance evidence. The additional evidence further strengthens the combined decision, and the system ultimately identifies the query as ‘CCVID_6‘. This case demonstrates the method’s ability to coordinate heterogeneous experts adaptively: subsequent model invocation is not fixed in advance, but is determined jointly by the current evidence and the query, enabling more reliable identity decisions in challenging conditions.

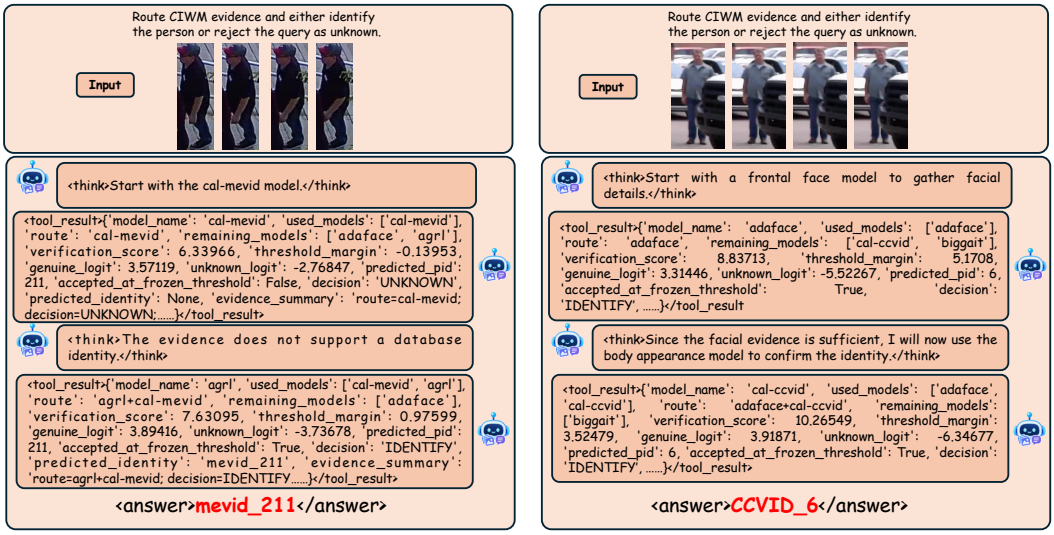

Figure 6: Additional CoT examples

A.3 Additional Reward Details.

In addition to the route-level reward, we use two auxiliary rewards to regularize the interaction protocol: a format reward and a tool-execution reward. These rewards evaluate whether the policy produces valid and executable interactions, but do not directly measure recognition accuracy.

Format Reward. The format reward checks every assistant turn in a generated trajectory. In the chain-of-thought mode, a valid turn must contain either a `<think>` block followed by one or more `<tool_call>` blocks, or a `<think>` block followed by an `<answer>` block. In the fast mode, a valid turn must contain either one or more `<tool_call>` blocks or a single `<answer>` block. In both modes, the XML-like structure must be well formed, all tags must be properly closed, and no additional top-level tags are allowed. Let $\mathbb{I}_{\text{fmt}}(t)$ indicate whether turn t satisfies these conditions. The trajectory-level format reward is the fraction of valid assistant turns:

$$R_{\text{format}} = \frac{1}{T}\sum_{t=1}^{T}\mathbb{I}_{\text{fmt}}(t), \tag{11}$$

where T is the number of assistant turns. A trajectory without an assistant turn receives a format reward of zero.

Tool Success Reward. The tool success reward evaluates the results returned by the tools. For each `<tool_result>` block, the returned content is parsed as a dictionary. A tool call is considered successful when the dictionary does not contain a non-empty `error` field. If N_{tool} denotes the number of tool results and N_{succ} denotes the number of successful results, the tool reward is

$$R_{\text{tool}} = \begin{cases} \dfrac{N_{\text{succ}}}{N_{\text{tool}}}, & N_{\text{tool}} > 0, \\ 0, & N_{\text{tool}} = 0. \end{cases} \tag{12}$$

This reward measures execution validity only; it does not assess whether the selected expert is optimal or whether the resulting identity decision is correct.

Accuracy Reward. To provide direct supervision for the agent's final recognition result, we introduce an answer-accuracy reward. For each generated trajectory, the reward function examines only the `<answer>` content in the final assistant turn and extracts the predicted identity string. The reward is set to 1 if this string exactly matches the ground-truth label and to 0 otherwise. A trajectory also receives a reward of 0 if it lacks a valid final turn, an `<answer>` tag, or parseable answer content. This reward evaluates only the correctness of the final identity output; it does not participate

in the computation of the APTC route utility, the CIEM certificate, or tool-execution status, and is therefore independent of the routing and interaction-protocol rewards. In the current GRPO configuration, the answer-accuracy reward is assigned a weight of 0.1 and is combined with the format reward, tool-execution reward, and APTC route reward to form the policy optimization objective.