A Quality-Guided Mixture of Score-Fusion Experts Framework for Human Recognition

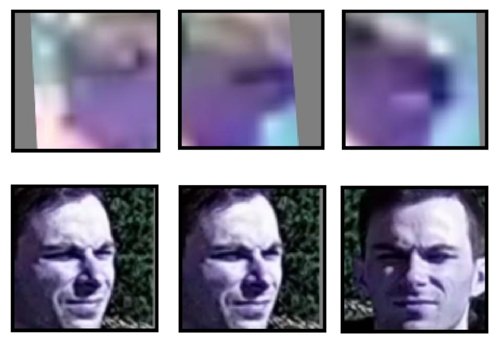
Jie Zhu[1], Yiyang Su[1], Minchul Kim[1], Anil Jain[1], and Xiaoming Liu[1]
[1]Michigan State University (MSU)

ICCV HONOLULU HAWAII OCT 19-23, 2025

Code

Motivation

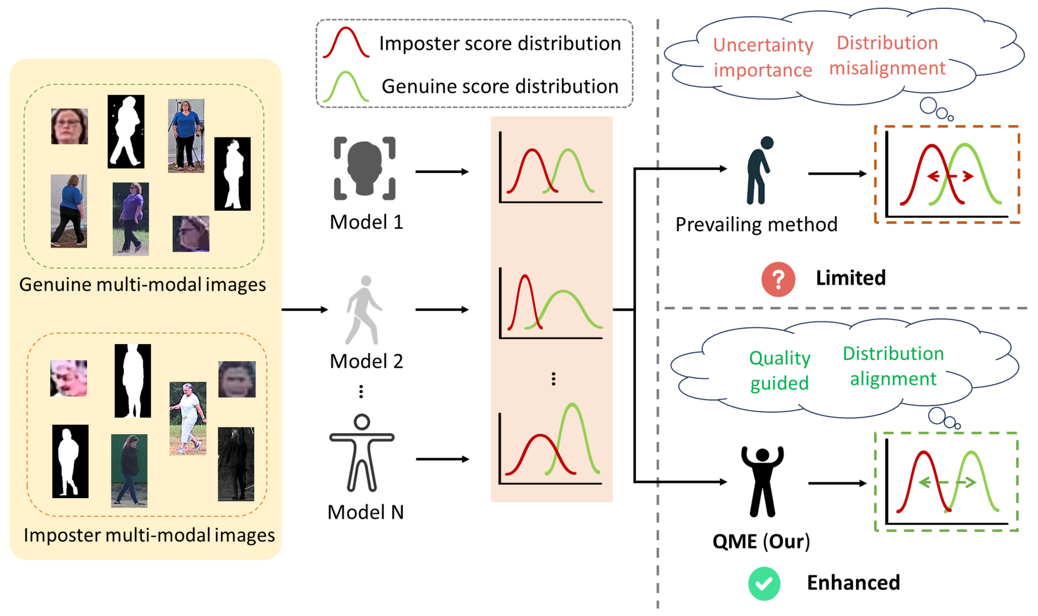
1. **Low quality vs high quality modality:**



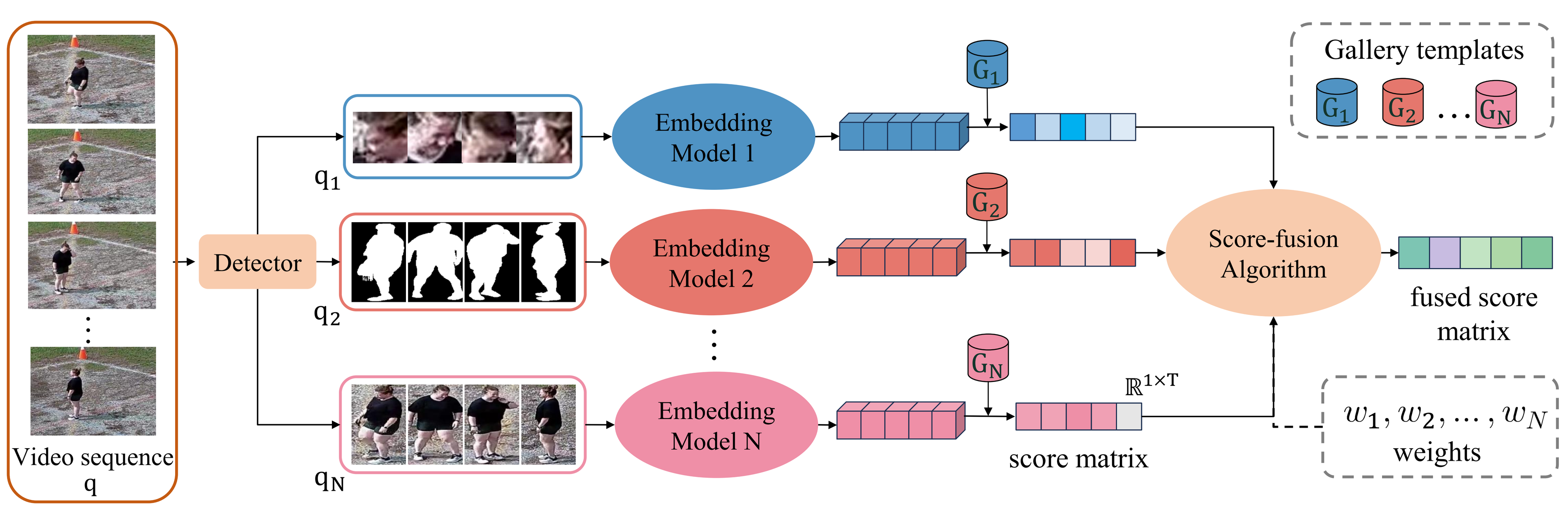
For a probe with **low-quality**, we should give a **lower weight**.

For a probe with **high-quality**, we should give a **higher weight**.

2. **Distribution misalignment & uncertainty importance:**

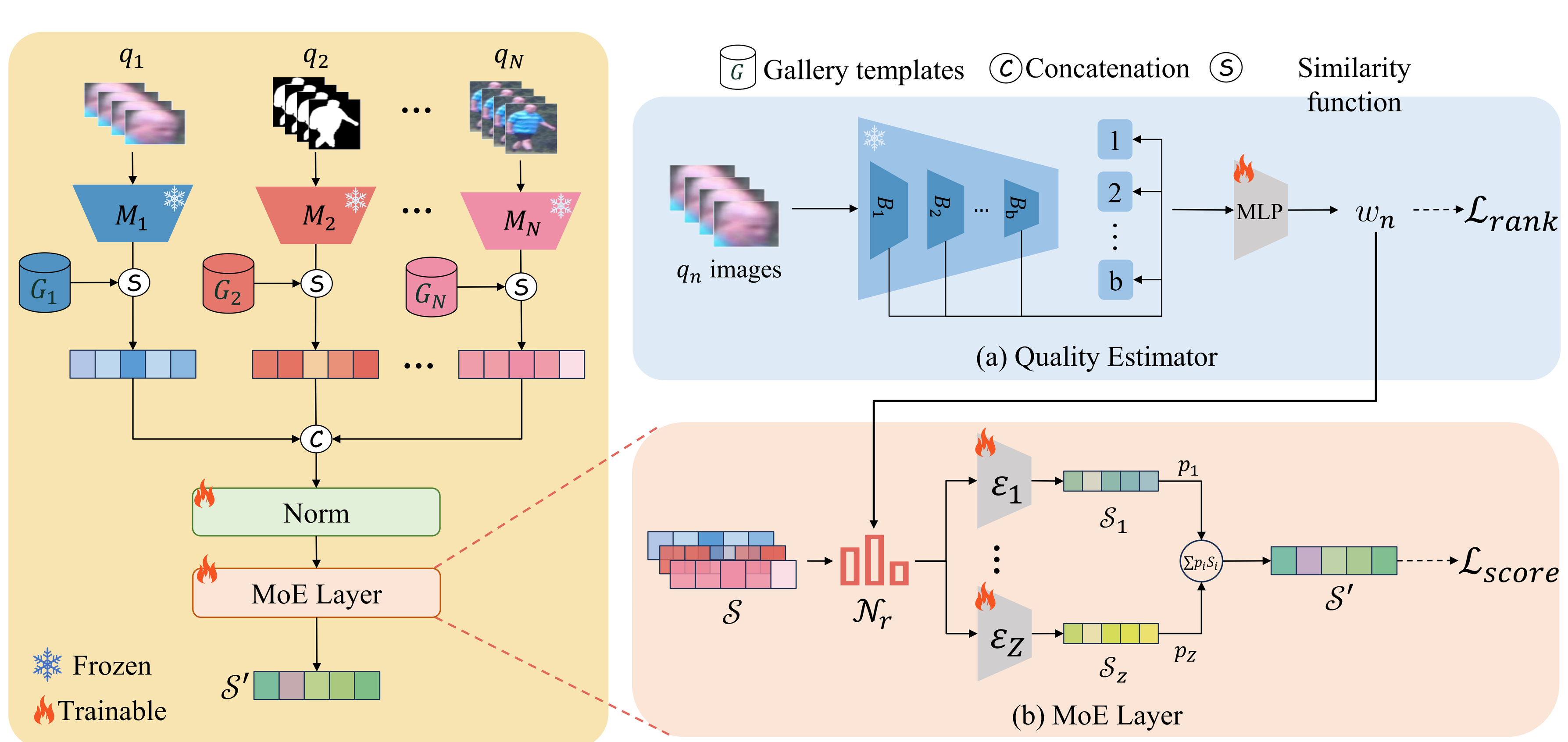


Systematic Whole-body Biometric Recognition



QME Framework

We propose a **Quality-Guided** Mixture of Score-Fusion Experts (QME) framework to fused the score matrices.



QME dynamically assigns weights to each expert based on **input quality**, encouraging **diverse specializations**.

Loss Functions

Pseudo Quality Loss

Provide pseudo quality labels based on the **ranking result** of the sample:

$$\mathcal{L}_{rank} = \sum_{i \in L} \mathrm{MSELoss}\left(w_i, \mathrm{ReLU}\left(\frac{\delta - r_i}{\delta - 1}\right)\right)$$

ranking result of the query feature

ranking threshold

Examples:
r_i=1, δ=10, Pseudo quality label: 1
r_i=5, δ=10, Pseudo quality label: 0.56
r_i=20, δ=10, Pseudo quality label: 0

Score Triplet Loss

Traditional triplet loss optimizes relative distances between samples:

$$\mathcal{L}_{tri} = \mathrm{ReLU}(d(a,p) - d(a,n) + m),$$

but it does not constrain the **values of negative samples**, which are crucial for calculating the **threshold** in verification and open-set search metrics.

To optimize these metrics, we introduce the score triplet loss:

$$\mathcal{L}_{score} = \mathrm{ReLU}(\mathcal{S}'_{nm}) + \mathrm{ReLU}(m - \mathcal{S}'_{mat})$$

Experiments

Our method consistently outperforms existing approaches across all four benchmarks, achieving SoTA performance in challenging metrics.

MEVID

Method	*Comb.*	Rank1↑	mAP↑	TAR↑	FNIR↓
*AdaFace** [22]	♦	25.0	8.1	5.4	98.8 ± 1.2
CAL [15]	♠	52.5	27.1	34.7	67.8 ± 7.3
AGRL [53]	■	51.9	25.5	30.7	69.4 ± 8.9
Z-score [47]		54.1	27.4	30.7	66.5 ± 7.0
Min-max [47]		52.8	24.7	25.0	71.3 ± 6.1
RHE [17]		52.8	24.8	25.3	71.2 ± 6.2
Weigthed-sum [35]	♦ ♠ ■	54.1	27.3	30.3	66.3 ± 7.0
Asym-AO1 [18]		52.5	22.9	23.6	71.7 ± 5.8
BSSF [49]		53.5	27.4	30.5	65.9 ± 7.2
Farsight [32]		53.8	25.4	26.6	69.8 ± 6.4
Ours (AdaFace-QE)		**55.7**	**28.2**	**32.9**	<u>64.6 ± 8.2</u>
Ours (CAL-QE)		<u>55.4</u>	<u>27.9</u>	<u>32.5</u>	$\mathbf{64.3 \pm 8.7}$

CCVID

Method	*Comb.*	Rank1↑	mAP↑	TAR↑	FNIR↓
*AdaFace** [22]	♦	94.0	87.9	75.7	13.0 ± 3.5
CAL [15]	♠	81.4	74.7	66.3	52.8 ± 13.3
*BigGait** [57]	♣	76.7	61.0	49.7	71.1 ± 6.1
Z-score [47]		92.2	90.6	73.9	15.1 ± 1.5
Min-max [47]		91.8	90.9	73.9	15.4 ± 2.5
RHE [17]		91.7	90.2	73.1	16.6 ± 2.5
Weigthed-sum [35]	♦ ♠ ♣	91.7	90.6	73.6	15.4 ± 1.8
Asym-AO1 [18]		92.3	90.0	74.0	15.9 ± 1.7
BSSF [49]		91.8	91.1	73.9	14.1 ± 1.3
Farsight [31]		92.0	<u>91.2</u>	73.9	13.9 ± 1.1
Ours (AdaFace-QE)		<u>92.6</u>	**91.6**	<u>75.0</u>	<u>13.3 ± 1.2</u>
Ours (CAL-QE)		**94.1**	90.8	**76.2**	$\mathbf{12.3 \pm 1.4}$

BRIAR

Method	*Comb.*	*Face Incl. Trt.*			*Face Restr. Trt.*		
		TAR↑	R20↑	FNIR↓	TAR↑	R20↑	FNIR↓
KPRPE [24]	♦	66.5	80.5	54.8	31.5	44.5	81.3
BigGait [57]	♣	66.3	93.1	72.7	61.0	90.4	76.3
CLIP3DReID [32]	♠	55.8	83.5	80.1	47.9	79.3	83.4
Farsight [31]		82.4	<u>95.8</u>	46.1	<u>65.7</u>	<u>91.0</u>	68.2
Ours		**84.5**	**96.0**	**41.2**	**67.9**	90.6	**64.1**

LTCC

Method	*Comb.*	Rank1↑	mAP↑	TAR↑	FNIR↓
*AdaFace** [22]	♦	18.5	5.9	2.4	99.8 ± 0.2
CAL [15]	♠	74.4	40.6	36.7	59.7 ± 7.3
AIM [56]	■	74.8	40.9	37.0	66.2 ± 9.2
Asym-AO1 [18]		71.2	32.9	19.1	76.3 ± 8.9
BSSF [49]		<u>73.5</u>	<u>39.1</u>	<u>34.2</u>	68.9 ± 8.5
Farsight [31]		73.2	37.8	31.3	72.4 ± 8.6
Ours		**73.8**	**39.6**	**35.0**	$\mathbf{64.3 \pm 8.0}$

Ablation Studies

(1) Effects of proposed component.

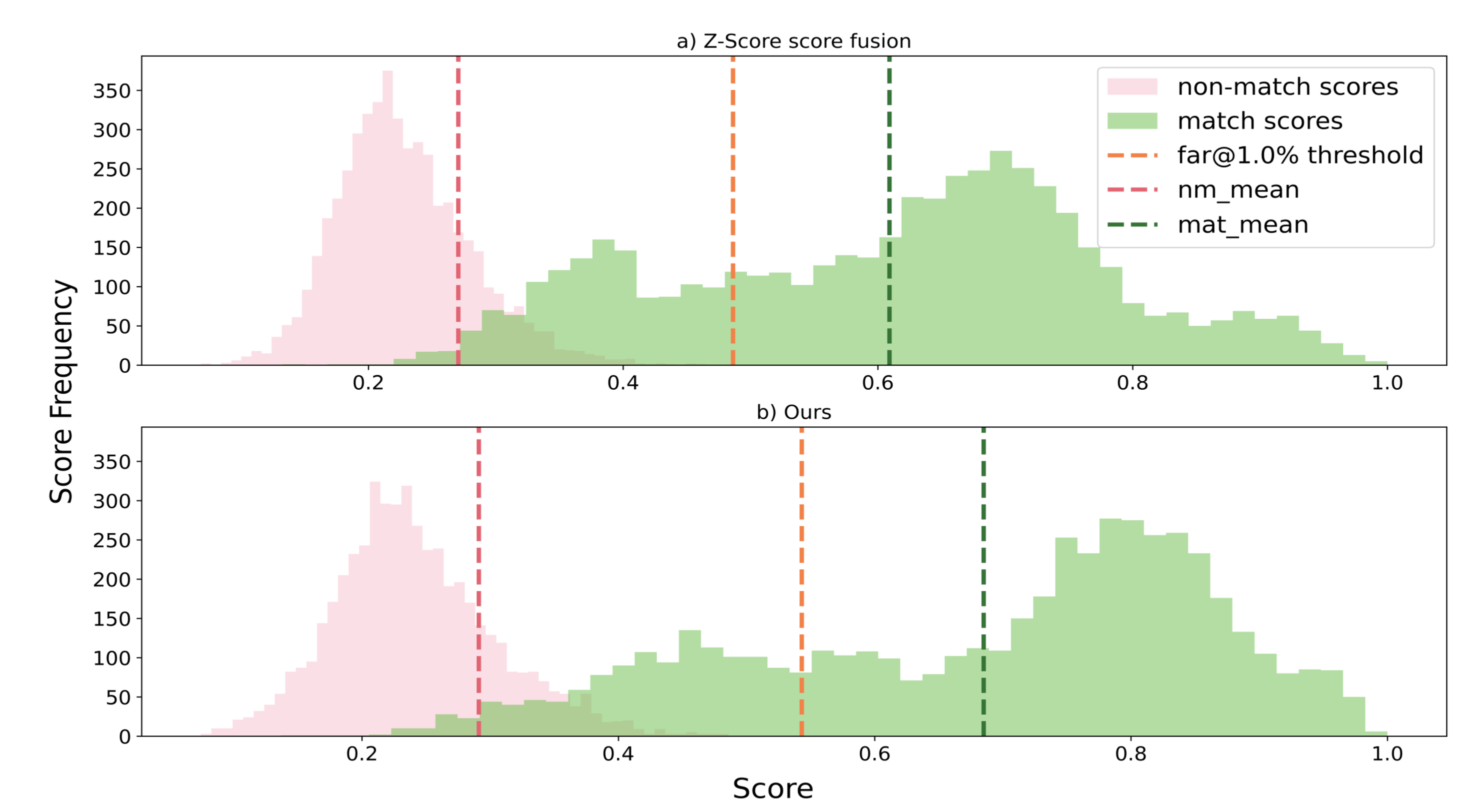
$\mathcal{L}_{score}$	QE	Z	Rank1↑	mAP↑	TAR↑	FNIR↓
✗	✗	1	49.4	21.6	23.3	84.0
✓	✗	1	53.8	24.5	25.3	70.4
✗	✗	2	54.1	25.5	30.8	65.4
✓	✗	2	55.1	27.0	31.3	66.5
✓	✓	2	**55.7**	**28.2**	**32.9**	**64.6**

(2) Effects of expert aggregation. Each expert specializes in **specific data scenarios and metrics**, and their aggregation leads to the best overall performance.

Expert	*Face Incl. Trt.*			*Face Restr. Trt.*		
	TAR↑	R20↑	FNIR↓	TAR↑	R20↑	FNIR↓
ε_1	<u>83.6</u>	<u>95.5</u>	<u>41.7</u>	62.0	**90.6**	<u>66.7</u>
ε_2	81.8	<u>95.5</u>	46.6	<u>65.0</u>	**90.6**	68.4
Ours ($\varepsilon_1 + \varepsilon_2$)	**84.5**	**95.7**	**41.2**	**67.9**	**90.6**	**64.1**

Visualizations

QME has a **clearer boundary** between match scores and non-match scores.



The predicted quality weight can **dynamically reflect** the quality of the input sample: