TransFIRA: Transfer Learning for Face Image Recognizability Assessment

Allen Tu[1,2], Kartik Narayan[3], Joshua Gleason[2], Jennifer Xu[2], Matthew Meyn[2], Tom Goldstein[1], Vishal Patel[3]

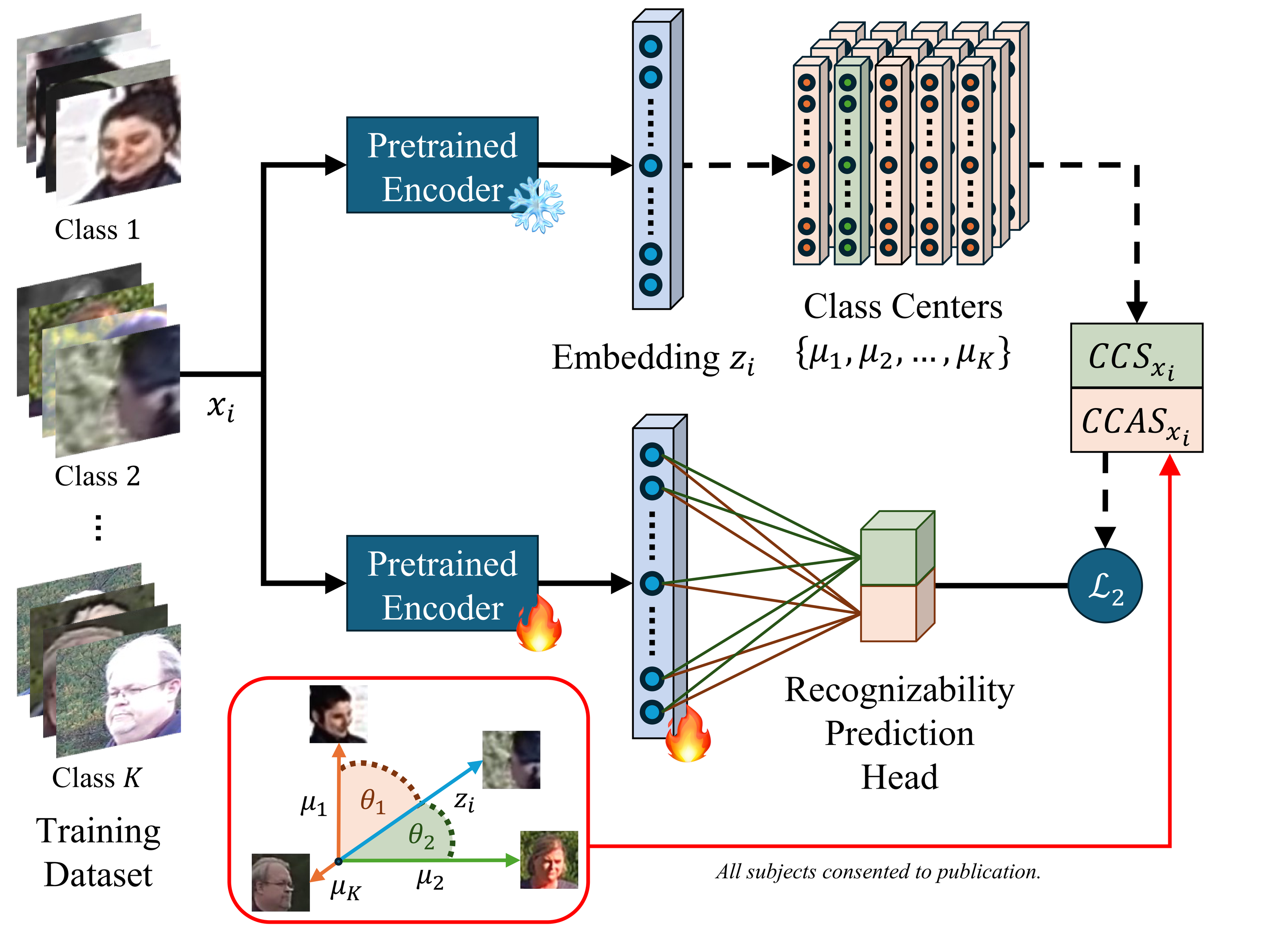
FG 2026 KYOTO 京

Oral Presentation

transfira.github.io

Motivation

How can we predict encoder-specific face recognizability from images and use it to guide template aggregation?



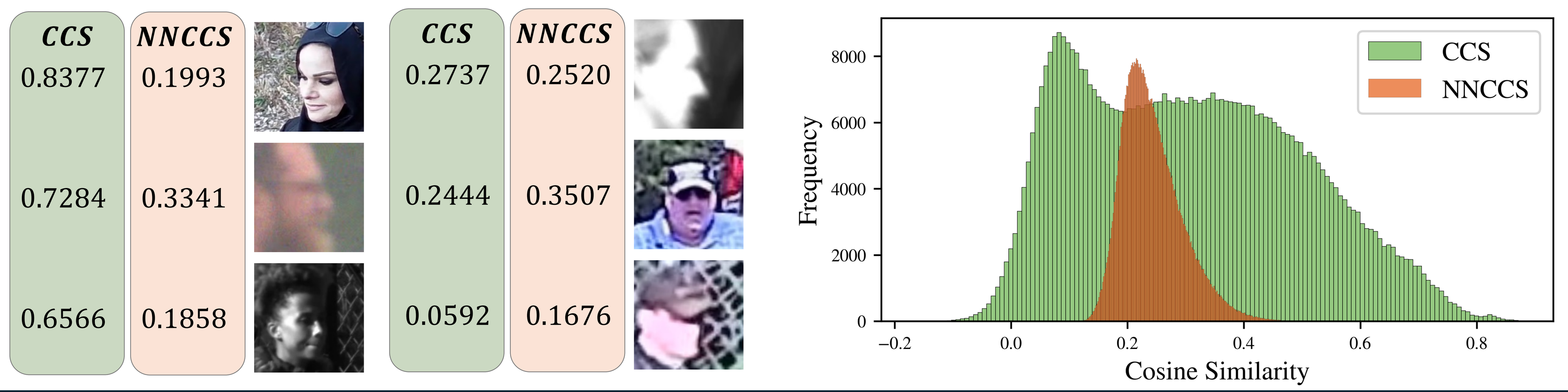
Recognizability: the likelihood that an image will be correctly identified by a given encoder.

Background

Template-based face recognition aggregates multiple images into a single representation, where even a few unrecognizable frames can corrupt it. Key benchmarks operate under **highly unconstrained conditions**:

- **BRIAR**: Video surveillance at altitude and range (private)
- **IJB-C**: Diverse web and video imagery (public)

Encoders must contend with extreme variation in pose, blur, illumination, and occlusion. Predictions from existing **face image quality assessment (FIQA)** metrics often fail to determine whether images are recognizable because they are detached from the **encoder's decision geometry**.



Method

TransFIRA adapts any pretrained encoder to predict recognizability from its own embedding geometry.

Our framework consists of three stages:

1. Image Recognizability

For each image x_i with embedding $z_i = \phi(x_i)$, we compute class-center similarities that measure its alignment with its identity.

- **Class Center Angular Similarity (CCS)** measures how closely an embedding aligns with the center of its own class:

$$CCS_{x_i} = \frac{z_i^\top \mu_{y_i}}{\|z_i\|_2 \, \|\mu_{y_i}\|_2}.$$

- **Nearest Nonmatch Class Center Angular Similarity (NNCCS)** measures its similarity to the most confusable imposter class:

$$NNCCS_{x_i} = \max_{j \neq y_i} \frac{z_i^\top \mu_j}{\|z_i\|_2 \, \|\mu_j\|_2}.$$

- **Class Center Angular Separation (CCAS)** is their difference:

$$CCAS_{x_i} = CCS_{x_i} - NNCCS_{x_i}.$$

$CCAS > 0$ indicates when z_i is closest to its own class center.

2. Recognizability Prediction Network

We extend the pretrained backbone with a **recognizability prediction head** to predict CCS and CCAS scores:

$$\hat{\mathbf{r}}_i = [\hat{CCS}_{x_i}, \hat{CCAS}_{x_i}]^\top = h_\psi(\phi(x_i)).$$

Fine-tuning both the backbone and head to predict the ground truth scores ensures **recognizability remains encoder-specific and fully aligned with the model's internal representation.**

3. Recognizability-Informed Template Aggregation

The predicted CCS and CCAS scores guide template aggregation:

- **Filtering:** Images with predicted $CCAS > 0$ are retained, ensuring only confidently recognized frames are included.
- **Weighting:** Embeddings are weighted by predicted CCS before averaging, emphasizing reliable samples near the class center.

Acknowledgements

This research is based upon work supported in part by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), via [2022-21102100005]. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of ODNI, IARPA, or the U.S. Government. The US Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright annotation therein.

Results

TransFIRA achieves **state-of-the-art results on BRIAR and IJB-C,** outperforming all prior FIQA methods.

Method	BRIAR Protocol 3.1 [5]: FNMR-ERC AUC							
	CosFace (BRIAR) [5,25]				ArcFace (WebFace) [9,27]			
	SC	10^{-3}	10^{-4}	10^{-6}	SC	10^{-3}	10^{-4}	10^{-6}
FaceQNet [10, 12]	0.1589	0.3957	0.5225	0.6903	0.1194	0.5037	0.7606	0.8924
MagFace [19]	0.1367	0.4314	0.5616	0.7336	0.0467	0.5426	0.8054	0.9200
SDD-FIQA [20]	0.0293	0.4513	0.5771	0.7335	0.0112	0.5527	0.8069	0.9148
CR-FIQA (S) [4]	0.1107	0.4199	0.5463	0.7093	0.0570	0.5409	0.7874	0.8975
CR-FIQA (L) [4]	0.4143	0.2708	0.3839	0.5640	0.3654	0.4387	0.6815	0.8386
DifFIQA (R) [1]	0.4416	0.2782	0.3820	0.5490	0.4657	0.4375	0.6622	0.8194
eDifFIQA (S) [2]	0.3610	0.2978	0.4064	0.5763	0.2577	0.4644	0.7035	0.8433
eDifFIQA (M) [2]	0.2870	0.3249	0.4267	0.5836	0.2265	0.4517	0.6895	0.8316
eDifFIQA (L) [2]	0.4377	0.2840	0.3855	0.5489	0.4415	0.4407	0.6658	0.8198
GraFIQs (S) [15]	0.0350	0.4363	0.5646	0.7344	-0.0194	0.5325	0.8102	0.9287
GraFIQs (L) [15]	0.1831	0.3881	0.5059	0.6732	0.1434	0.5074	0.7643	0.8914
CLIB-FIQA [21]	0.5004	0.2536	0.3591	0.5340	0.4404	0.4508	0.6758	0.8281
Ours (CCAS)	**0.8230**	**0.1558**	**0.2734**	**0.4752**	**0.6568**	**0.4071**	**0.6335**	**0.8068**
Ours (CCS)	**0.8566**	**0.1536**	**0.2722**	**0.4753**	**0.6401**	**0.3934**	**0.6315**	**0.8076**

InsightFace ArcFace verification on BRIAR improves **from 0.4269 to 0.8715 TAR at 10^{-3} FMR,** despite ArcFace never being trained on BRIAR.

Method	BRIAR Protocol 3.1 [5]: TAR at Fixed FMR					
	CosFace (BRIAR) [5,25]			ArcFace (WebFace) [9,27]		
	10^{-3}	10^{-4}	10^{-6}	10^{-3}	10^{-4}	10^{-6}
Average (Baseline)	0.9077	0.8423	0.7500	0.4269	0.1923	0.0846
FaceQNet [10, 12]	0.9077	0.8500	0.7577	0.4385	0.2385	0.0923
MagFace [19]	0.9077	0.8423	0.7538	0.4385	0.2154	0.0923
SDD-FIQA [20]	0.9038	0.8462	0.7538	0.4269	0.2154	0.0923
CR-FIQA (S) [4]	0.9038	0.8462	0.7577	0.4538	0.2577	0.1077
CR-FIQA (L) [4]	0.9231	0.8731	0.8077	0.5346	0.3231	0.1654
DifFIQA (R) [1]	0.9423	0.8769	0.8269	0.6192	0.4885	0.3885
eDifFIQA (S) [2]	0.8885	0.8462	0.7962	0.5500	0.3731	0.2808
eDifFIQA (M) [2]	0.6769	0.5923	0.4423	0.3962	0.3000	0.2115
eDifFIQA (L) [2]	0.8500	0.8000	0.7154	0.6154	0.4692	0.3577
GraFIQs (S) [15]	0.9077	0.8423	0.7500	0.4269	0.1923	0.0846
GraFIQs (L) [15]	0.9077	0.8423	0.7500	0.4269	0.1923	0.0846
CLIB-FIQA [21]	0.9154	0.8654	0.7885	0.5038	0.3115	0.1769
Ours (CCAS Filter)	**0.9846**	**0.9231**	**0.8846**	**0.8715**	**0.7510**	**0.5703**
Ours (CCS Weight)	0.9615	0.9038	0.8462	0.6269	0.4308	0.2692
Ours (Filter + Weight)	**0.9885**	**0.9269**	**0.8885**	**0.8715**	**0.7510**	**0.5984**

We extend **TransFIRA** to **body recognizability prediction,** improving template aggregation on BRIAR.

Method	TAR@1e-3	TAR@1e-4	TAR@1e-6
Average (Baseline)	0.5934	0.3278	0.0851
Calibrated CCAS Filter	0.6037	0.3485	0.0933
Calibrated CCAS Weight	**0.6452**	**0.3693**	0.1037
Calibrated Filter + Weight	0.6328	0.3568	**0.1079**

See our paper for IJB-C results and explainability analyses.