UNIVERSITÄT ZÜRICH
UNIVERSIDADE FEDERAL DE MINAS GERAIS

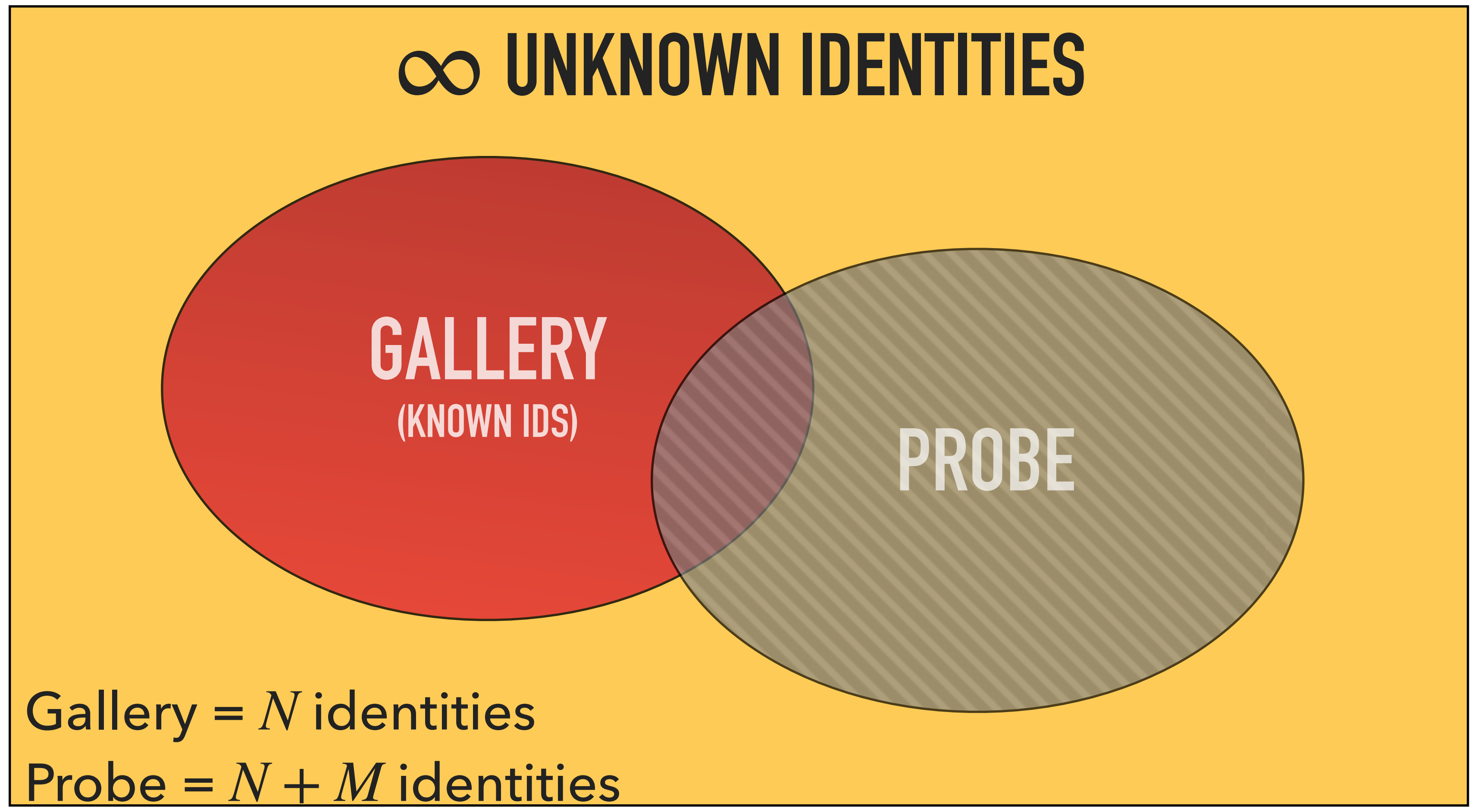
SIBGRAPI 2023
36TH CONFERENCE ON GRAPHICS, PATTERNS AND IMAGES

Authors:
Rafael H. Vareto (UFMG)
Manuel Günther (UZH)
William R. Schwartz (UFMG)

OPEN-SET FACE RECOGNITION WITH NEURAL ENSEMBLE, MAXIMAL ENTROPY LOSS AND FEATURE AUGMENTATION

OPEN-SET FACE RECOGNITION

Recognize faces of subjects that may have not been seen during training.



DOMINANT CHALLENGES

Face Recognition Task:

- Imbalanced datasets (bias)
- Classes and samples disparity (overfitting)
- Different face domains (pose, occlusion, quality)

Open Set Task:

- Neural networks do not know the unknown
- Unknown persons may resemble known subjects
- *Watchlists*: operate at low false-positive rate

RELATED WORKS

Towards Open-Set Face Recognition using Hashing Functions

Rafael Vareto, Samira Silva, Filipe Costa, William Robson Schwartz
Smart Surveillance Interest Group, Department of Computer Science
Universidade Federal de Minas Gerais, Minas Gerais, Brazil
{rafaelvareto, samirasilva, fcosta, william}@dcc.ufmg.br

Reducing Network Agnostophobia

Akshay Raj Dhamija, Manuel Günther, and Terrance E. Boult
Vision and Security Technology Lab, University of Colorado Colorado Springs
{adhamija | mgunther | tboult} @ vast.uccs.edu

Watchlist Adaptation: Protecting the Innocent

Manuel Günther
Department of Informatics
University of Zurich
Zürich, Switzerland
guenther@ifi.uzh.ch

Akshay Raj Dhamija, Terrance E. Boult
Vision and Security Technology
University of Colorado Colorado Springs
Colorado Springs, USA
{dhamija,tboult}@vast.uccs.edu

Manifold Mixup: Better Representations by Interpolating Hidden States

Vikas Verma [*12] Alex Lamb [*2] Christopher Beckham [2] Amir Najafi [3] Ioannis Mitliagkas [2] David Lopez-Paz [4] Yoshua Bengio [2]

senselab

PROPOSED METHOD: MAXIMAL ENTROPY LOSS (MEL)

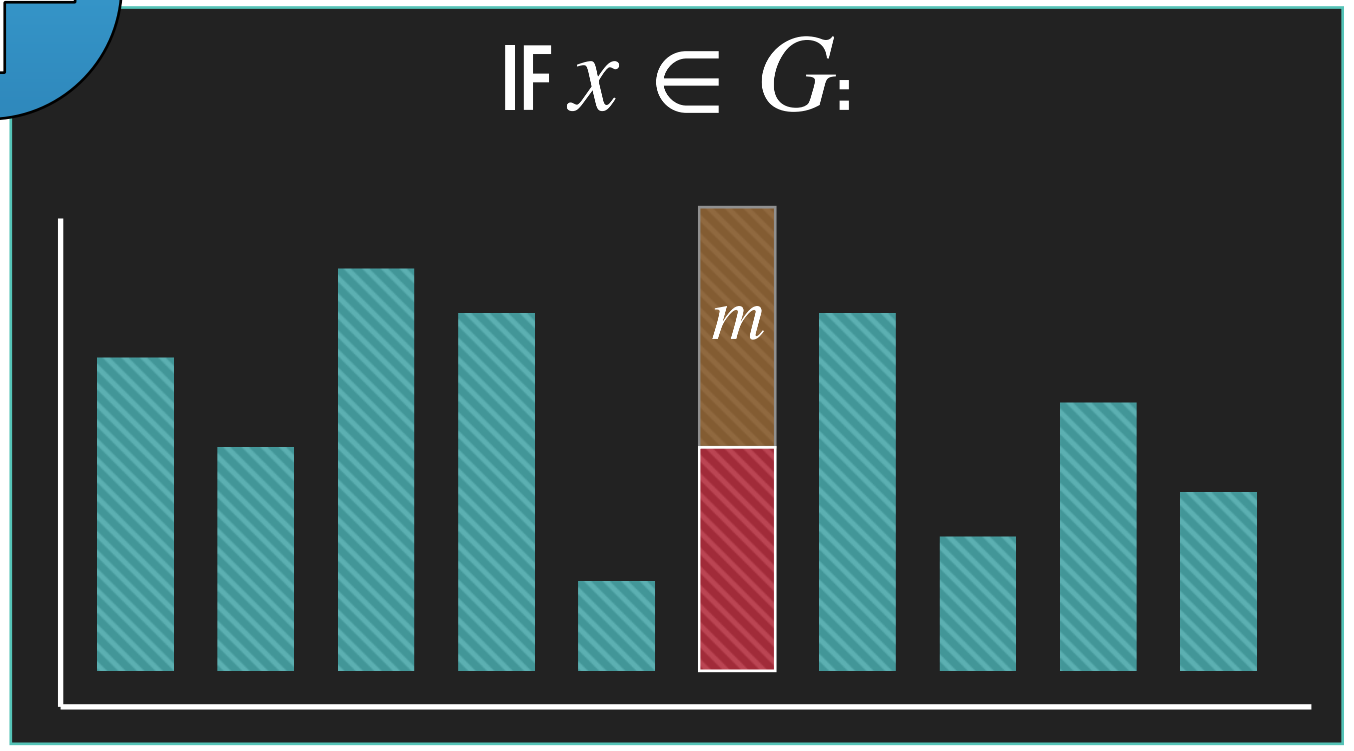
ISSUE: Stop network from returning high scores for unknown samples

$$J_M = \begin{cases} -\log A_{Sm}(x) & \text{if } x \in G \\ -\frac{1}{|G|} \sum_{g=1}^{|G|} \log A_S(x) & \text{if } x \notin G \end{cases}$$

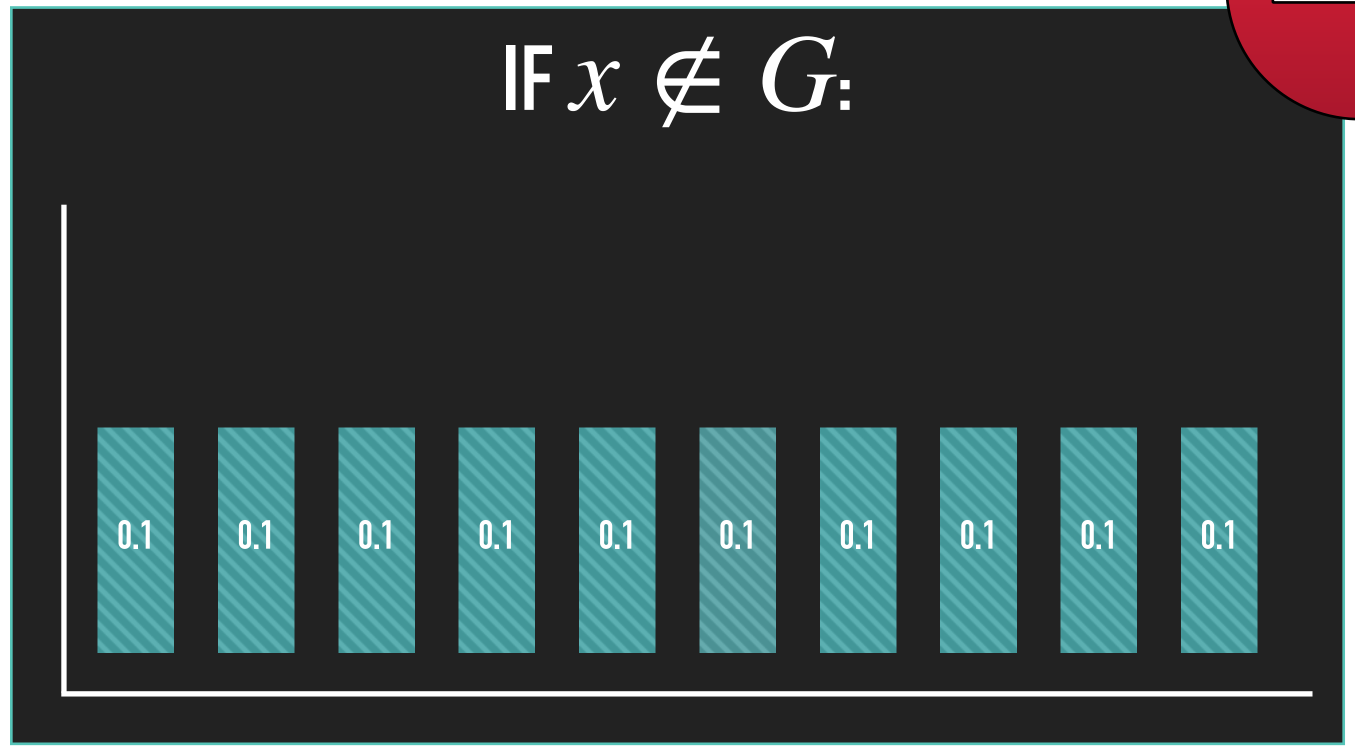
$$A_{Sm}(l_g) = \frac{e^{l_g(x)-m}}{e^{l_g(x)-m} + \sum_{\hat{g} \neq g}^{G} e^{l_{\hat{g}}(x)}}$$

$$A_S(l_g) = \frac{e^{l_g(x)}}{\sum_{\hat{g}}^{G} e^{l_{\hat{g}}(x)}}$$

Gallery/Positive Samples



Non-Gallery/Negative Samples



senselab

PROPOSED METHOD: OPTIMIZED MIXUP AUGMENTATION (OMU)

ISSUE: Negative samples from different distributions do not cooperate

$$\bar{z} = \lambda \cdot z_i \; + \; (1 - \lambda) \cdot z_j$$

$$\text{s.t. } z_j = \text{argmax}_{(z_{i'}, g_{i'}) \in G} \; \cos(z_i, z_{i'}) \;\; \wedge \;\; g_i \neq g_{i'}$$

λ

0.3	0.7	0.2	0.9	0.5	0.6	0.8	0.5	0.3	0.4

$+(1 - \lambda)$

0.5	0.4	0.7	0.9	0.4	0.2	0.8	0.3	0.5	0.1

$\lambda = 0.5$

0.40	0.55	0.45	0.90	0.45	0.40	0.80	0.40	0.40	0.25

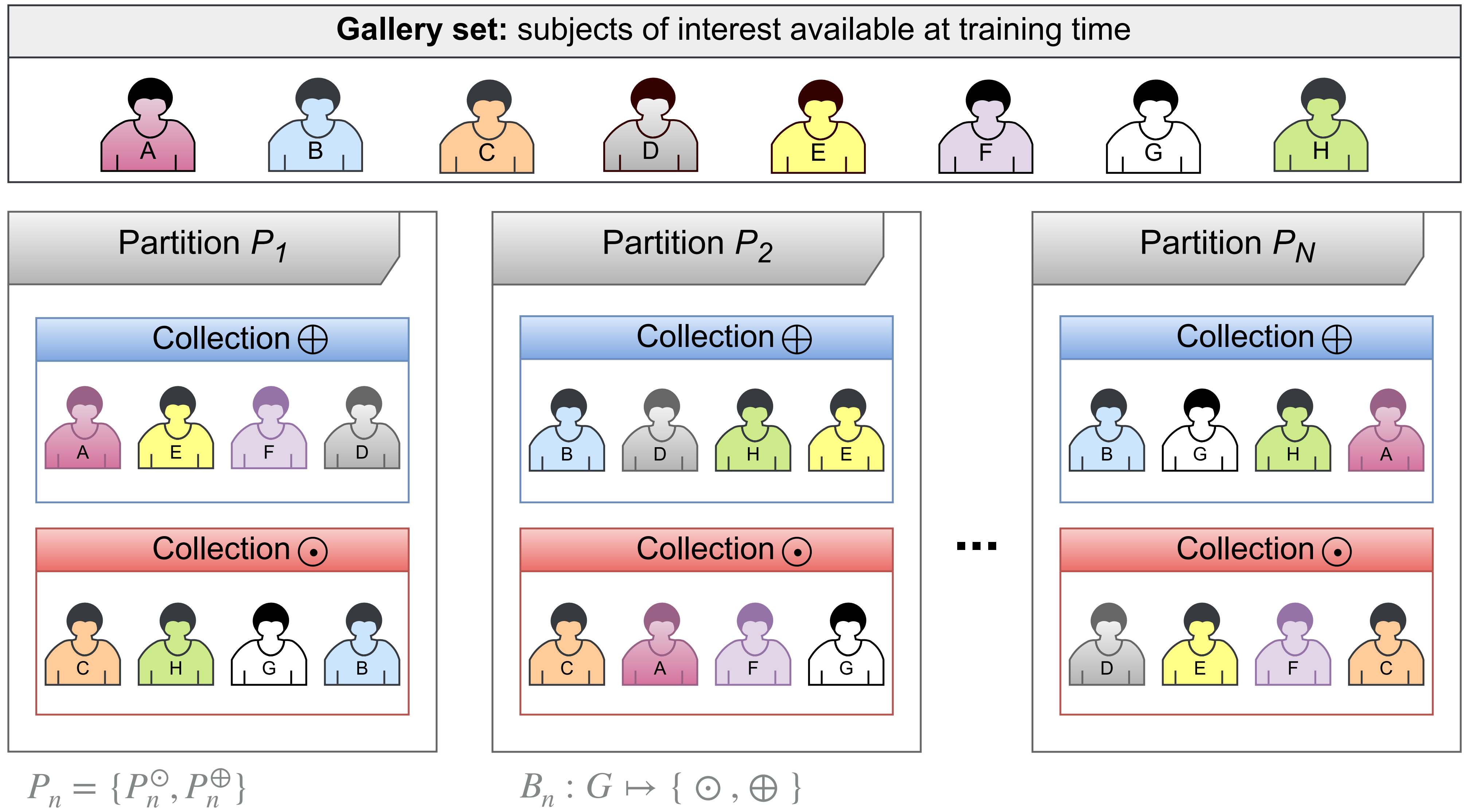
PROPOSED METHOD: NEURAL ADAPTER ENSEMBLE (NAE)

ISSUE: Imbalanced datasets bias neural networks toward majority classes

NAE data partitioning:

- Split the gallery set into N disjoint partitions $P_n = \{P_n^{\odot}, P_n^{\oplus}\}$
- Assign temporary new labels: $\oplus$ $\odot$
- Allocate identity $g \in G$ into $P_n^{\odot}$ or $P_n^{\oplus}$ following function $B_n : G \mapsto \{ \odot , \oplus \}$
- Fit classifier C_n to corresponding binary split P_n

Adapter from "Unconstrained face identification using ensembles trained on clustered data"



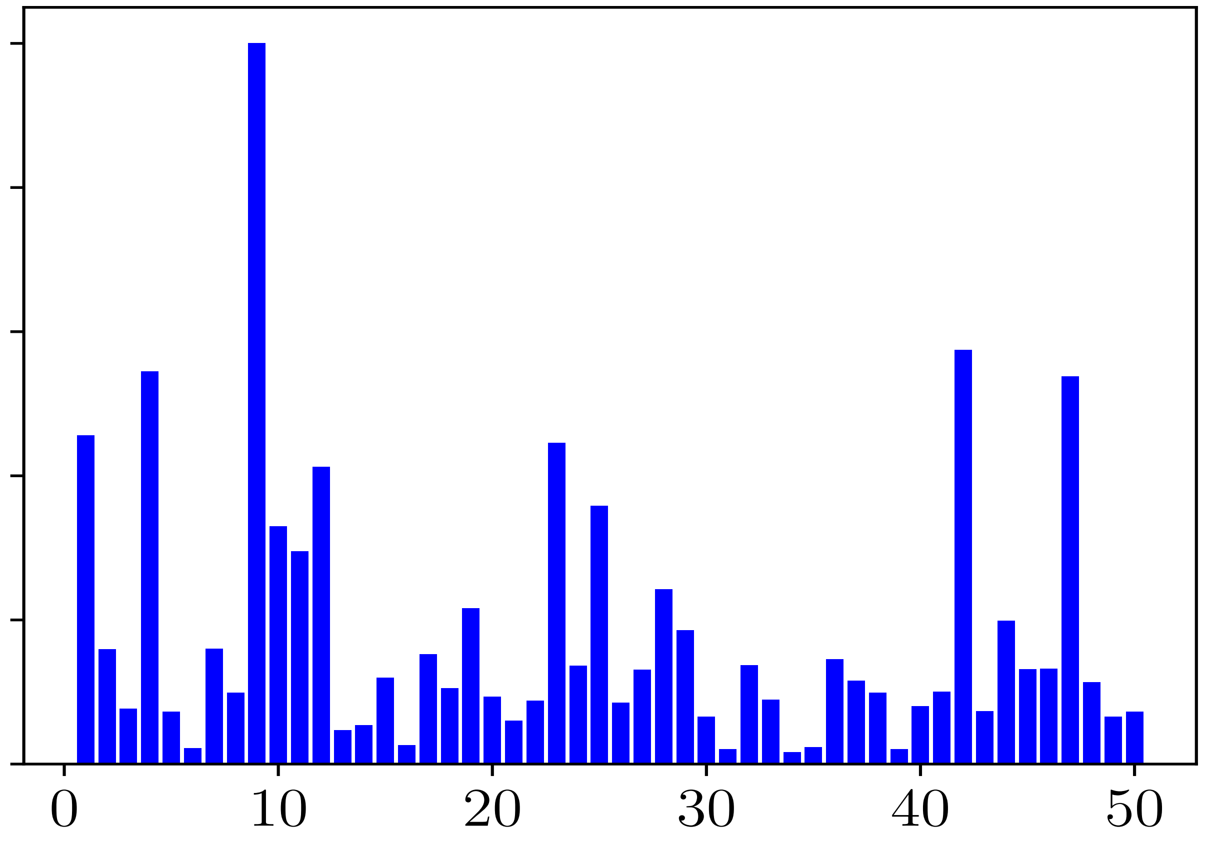
PROPOSED METHOD: NEURAL ADAPTER ENSEMBLE (NAE)

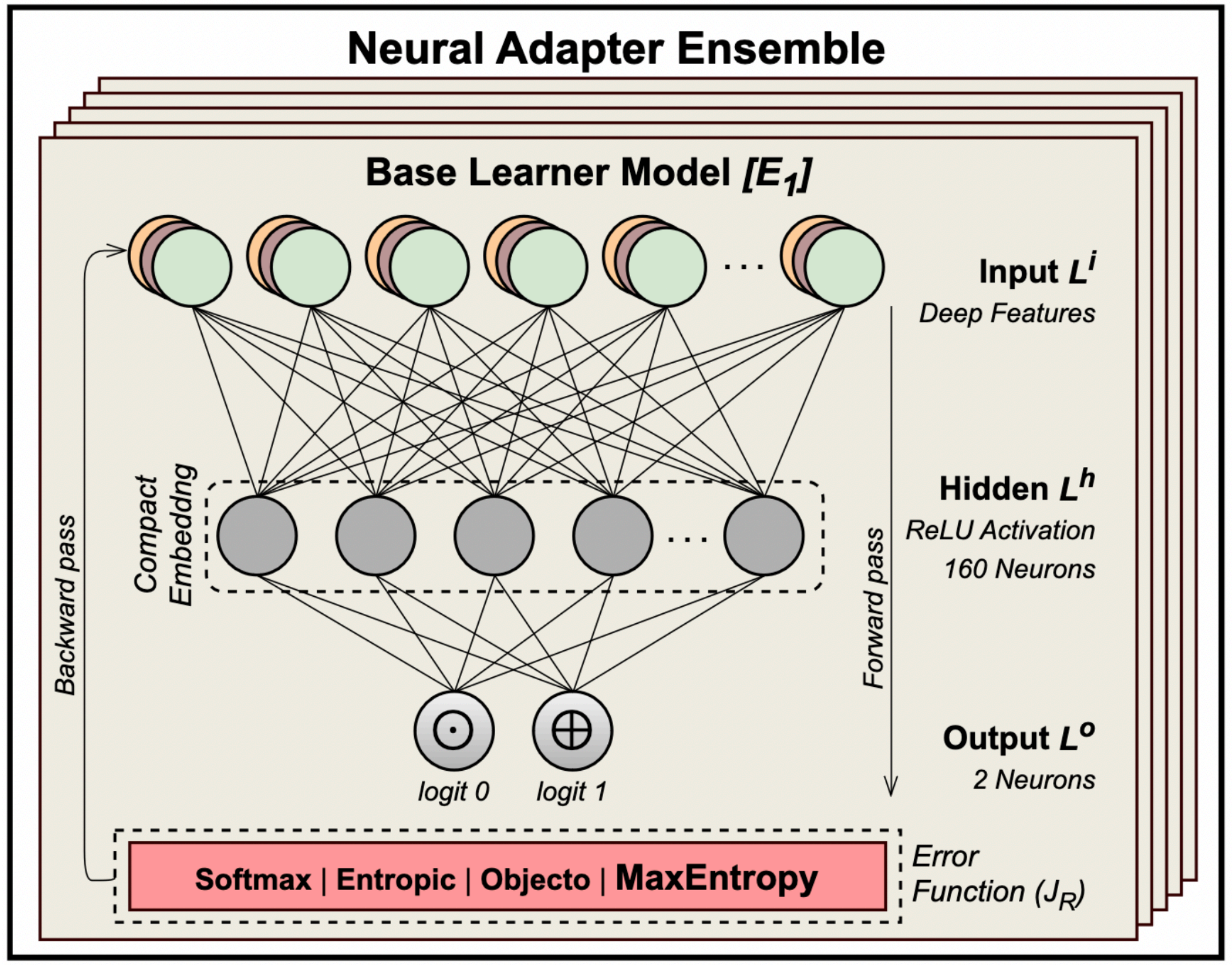
ISSUE: Map binary ids { ⊙ , ⊕ } back to original gallery ids

Rank of candidates:

$$\mathrm{sim}(z_p, g) = \sum_n [L_n^o]^{B_n(g)}(z_p),\ \forall\ g \in G$$

$P_n = \{P_n^{\odot}, P_n^{\oplus}\}$ $\quad$ $B_n : G \mapsto \{ \odot , \oplus \}$





DATASETS: LFW AND IJB-C

Labeled Faces in the Wild

- 13.233 near-frontal images from approximately six thousand persons
- Distinct poses, expressions, scenes and lighting

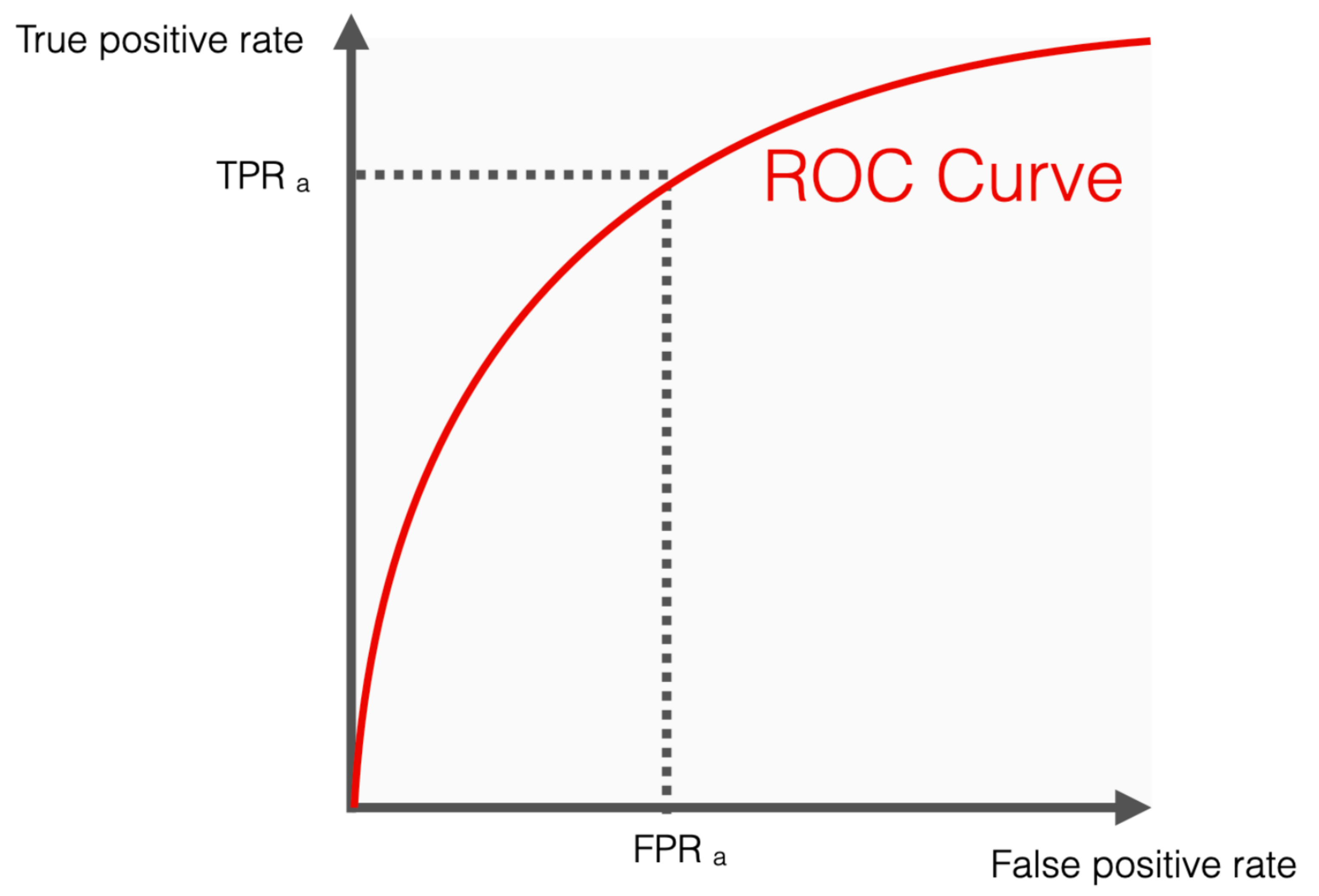
IARPA Janus Benchmark C

- Two gallery-sets (G1/G2) holding around 1.8K identities each
- Profile views, occlusion and low-resolution probe images

senselab

METRIC: OPEN–SET RECEIVER OPERATING CHARACTERISTIC

- Belongs to the gallery and retrieves correct ID
- Vertical and Horizontal axes:
  - True-positive Identification Rate
  - False-positive Identification Rate
- Hyper-parameters:
  - NAE size $|E|$, MEL margin m, OMU factor λ



LFW EVALUATION. OPEN-SET ASSESSMENT TO SELECT OPTIMAL VALUES FOR PARAMETERS λ, h AND $|E|$.

Parameters	$\|E\|$					m					λ				
DIR/VALUES	0.10	0.30	0.50	0.75	1.00	0.10	0.20	0.30	0.40	0.50	0.55	0.65	0.75	0.85	0.95
TPIR@FPIR = 1.00	0.82	0.92	**0.94**	**0.94**	**0.94**	**0.94**	**0.94**	**0.94**	**0.94**	0.93	0.94	0.94	0.94	**0.95**	0.94
TPIR@FPIR = 0.10	0.71	0.85	0.86	**0.88**	**0.88**	0.86	0.86	**0.87**	**0.87**	**0.87**	0.87	0.87	**0.89**	**0.89**	0.87
TPIR@FPIR = 0.01	0.57	0.72	0.73	**0.75**	**0.75**	0.73	0.74	**0.77**	0.76	0.74	**0.77**	**0.77**	**0.77**	0.74	0.75

EVALUATION: FEATURE AUGMENTATIONS STRATEGIES ON NAN

AUGMENTATION ANALYSIS. EVALUATION OF NAN TRAINED WITH CEL OR MEL ASSOCIATED WITH DIFFERENT AUGMENTATION SCHEMES ON IJB-C.

Method	Detection and Identification Rate (TPIR@) FPIR=1	FPIR=0.1	FPIR=0.01	FPIR=0.001
CEL	0.44	0.23	0.09	0.03
MEL+LFW	0.58	0.23	0.10	0.03
MEL+SFA	**0.68**	**0.53**	0.31	0.05
MEL+MMU	**0.68**	0.52	0.28	0.04
MEL+OMU	0.66	0.51	**0.33**	**0.10**

CEL: Cross-Entropy Loss
MEL: Maximal Entropy Loss

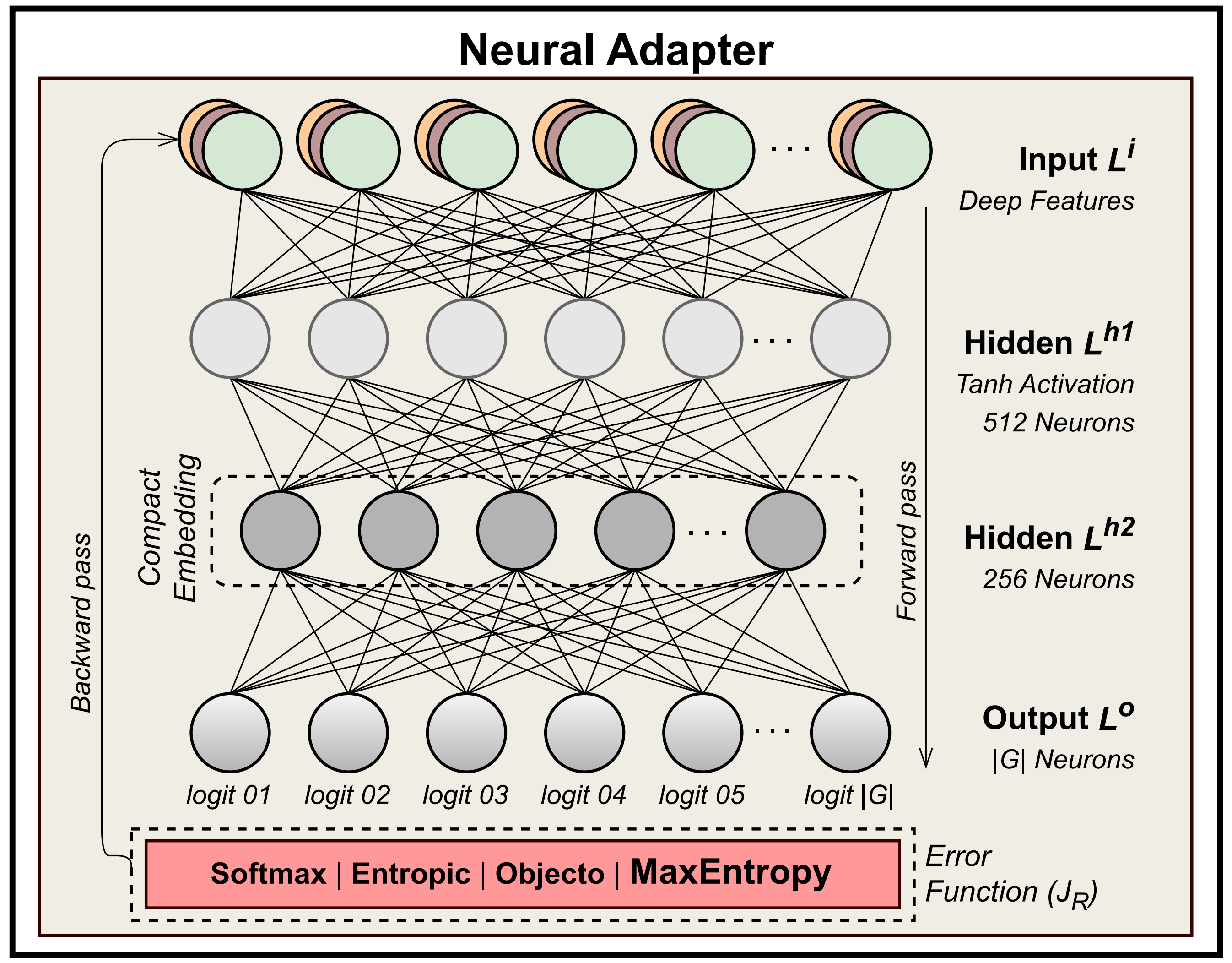
SFA: Stochastic Feature Augmentation
MMU: Manifold Mix-Up Augmentation
OMU: Optimized Mix-Up Augmentation

LFW: Labelled Faces in the Wild dataset

Watchlist Adaptation: Protecting the Innocent

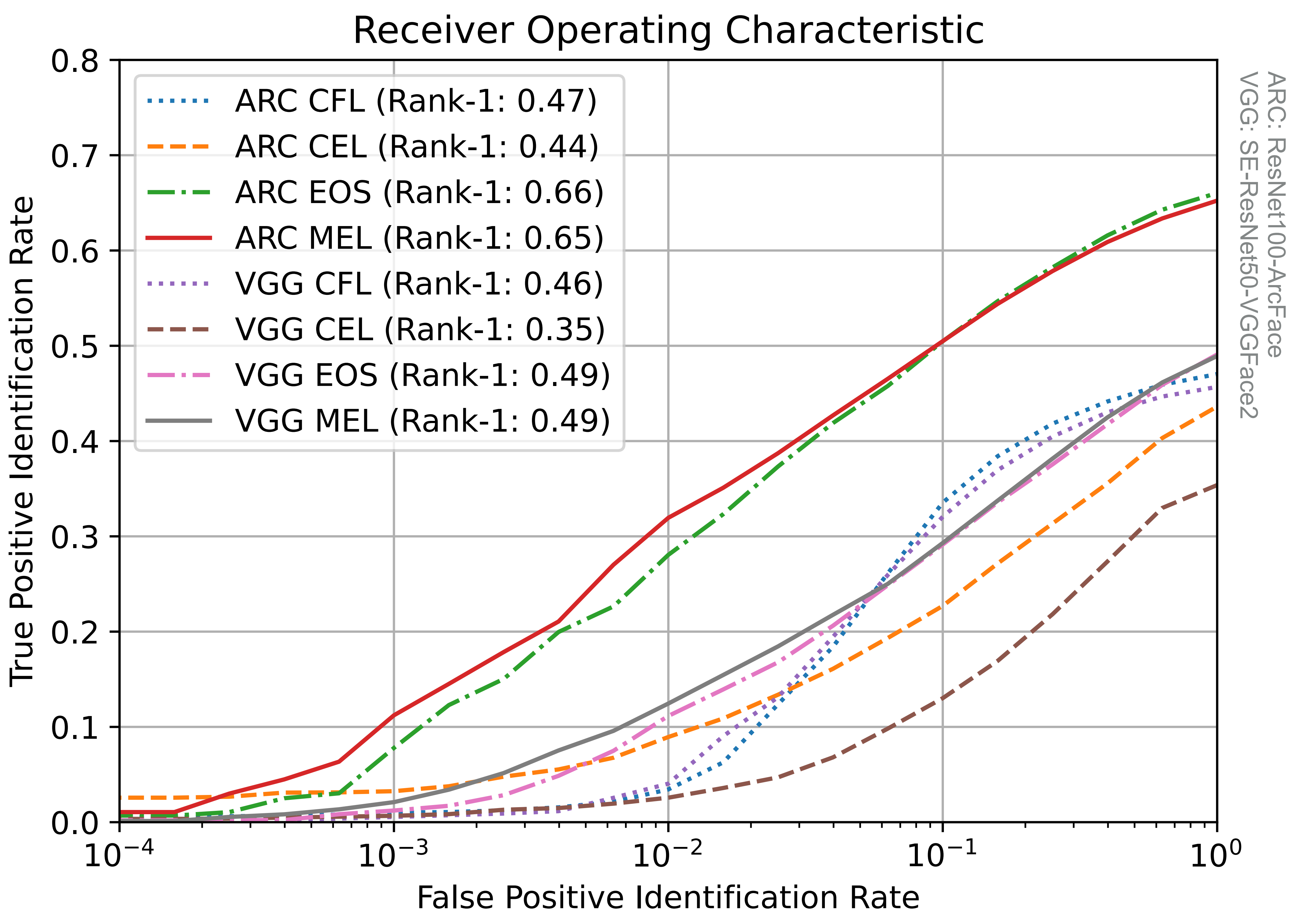
Manuel Günther
Department of Informatics
University of Zurich
Zürich, Switzerland
guenther@ifi.uzh.ch

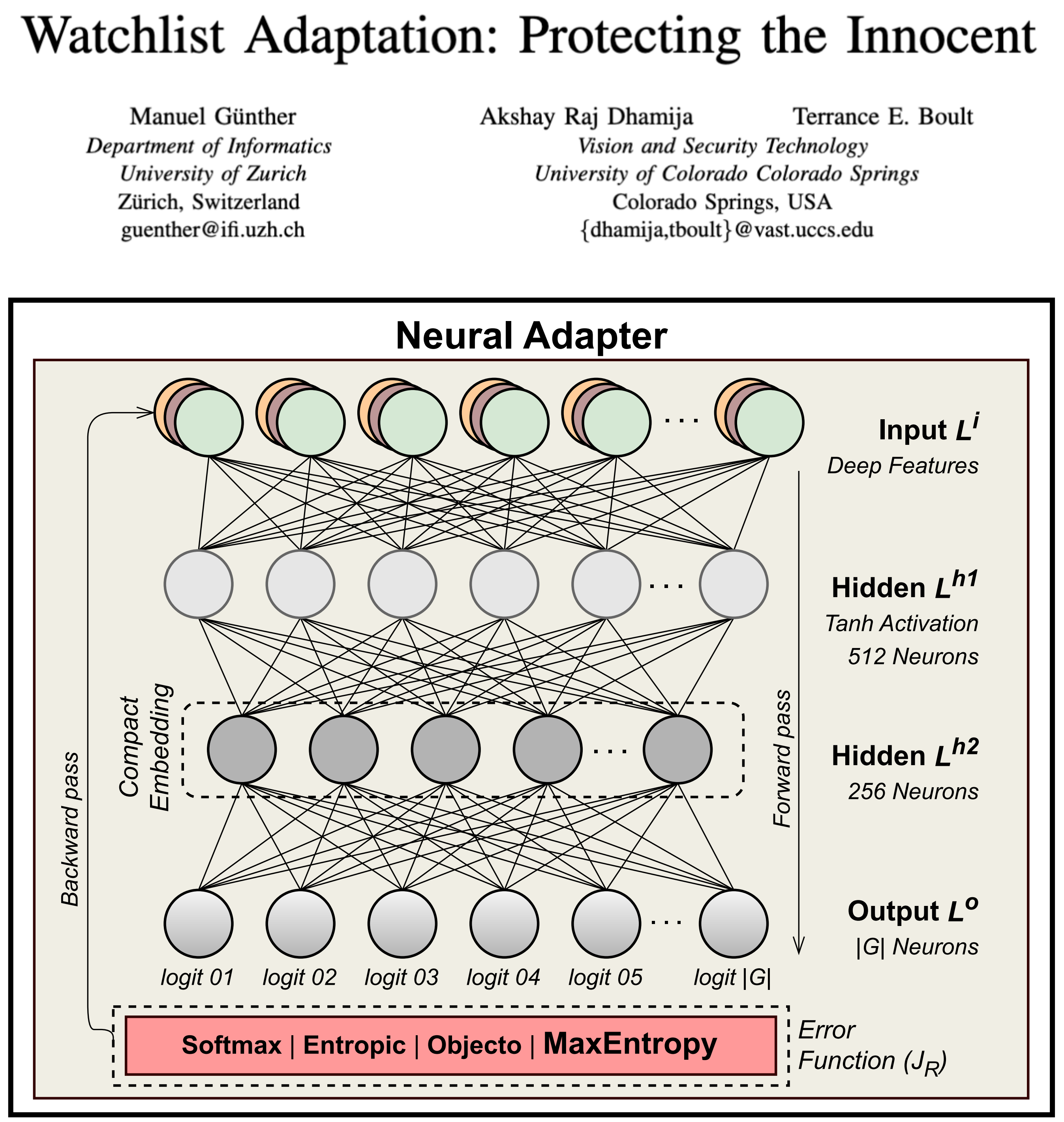
Akshay Raj Dhamija, Terrance E. Boult
Vision and Security Technology
University of Colorado Colorado Springs
Colorado Springs, USA
{dhamija,tboult}@vast.uccs.edu



IARPA Janus Benchmark C

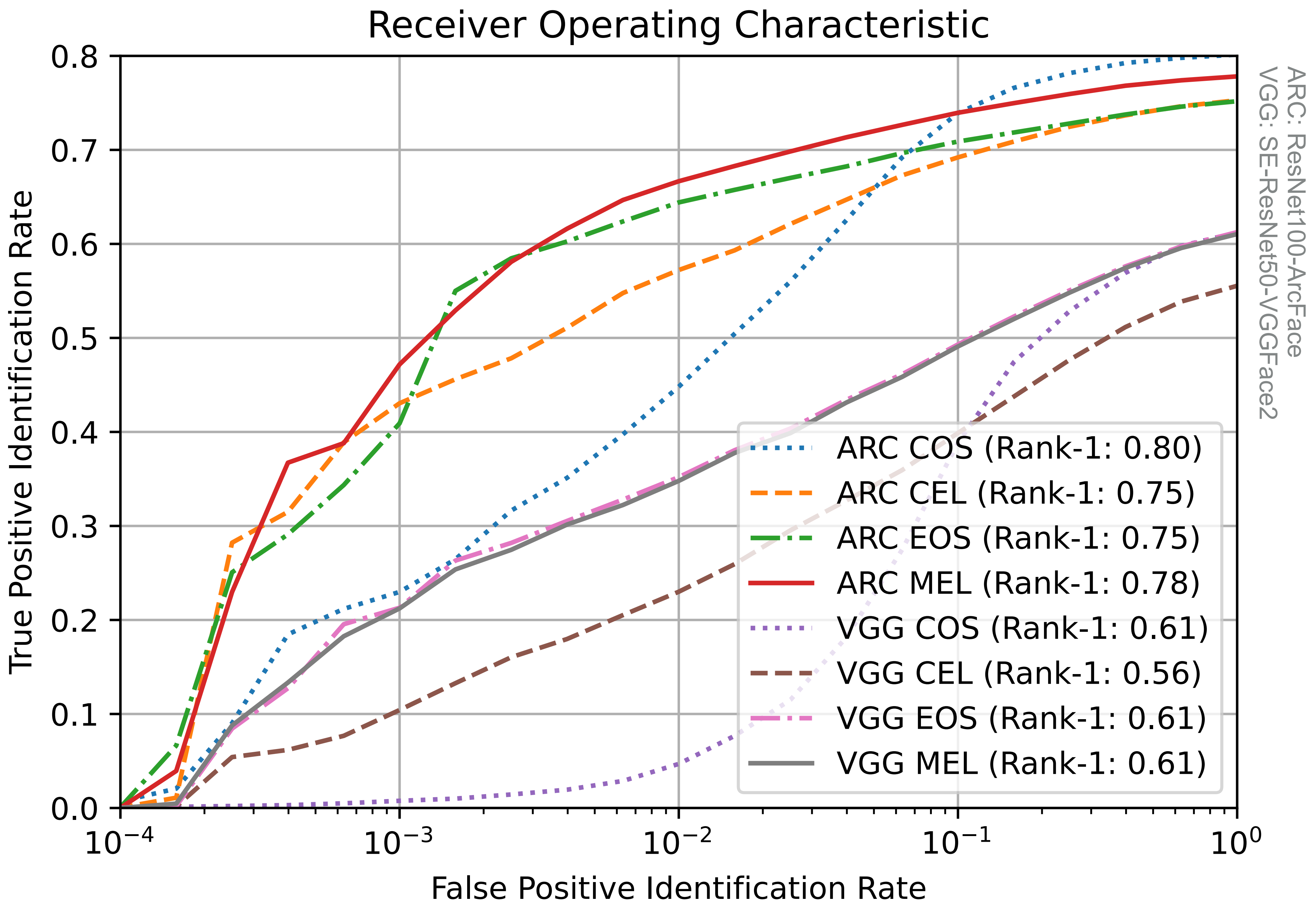
EVALUATION: CONTRASTING LOSS FUNCTIONS ON NAN

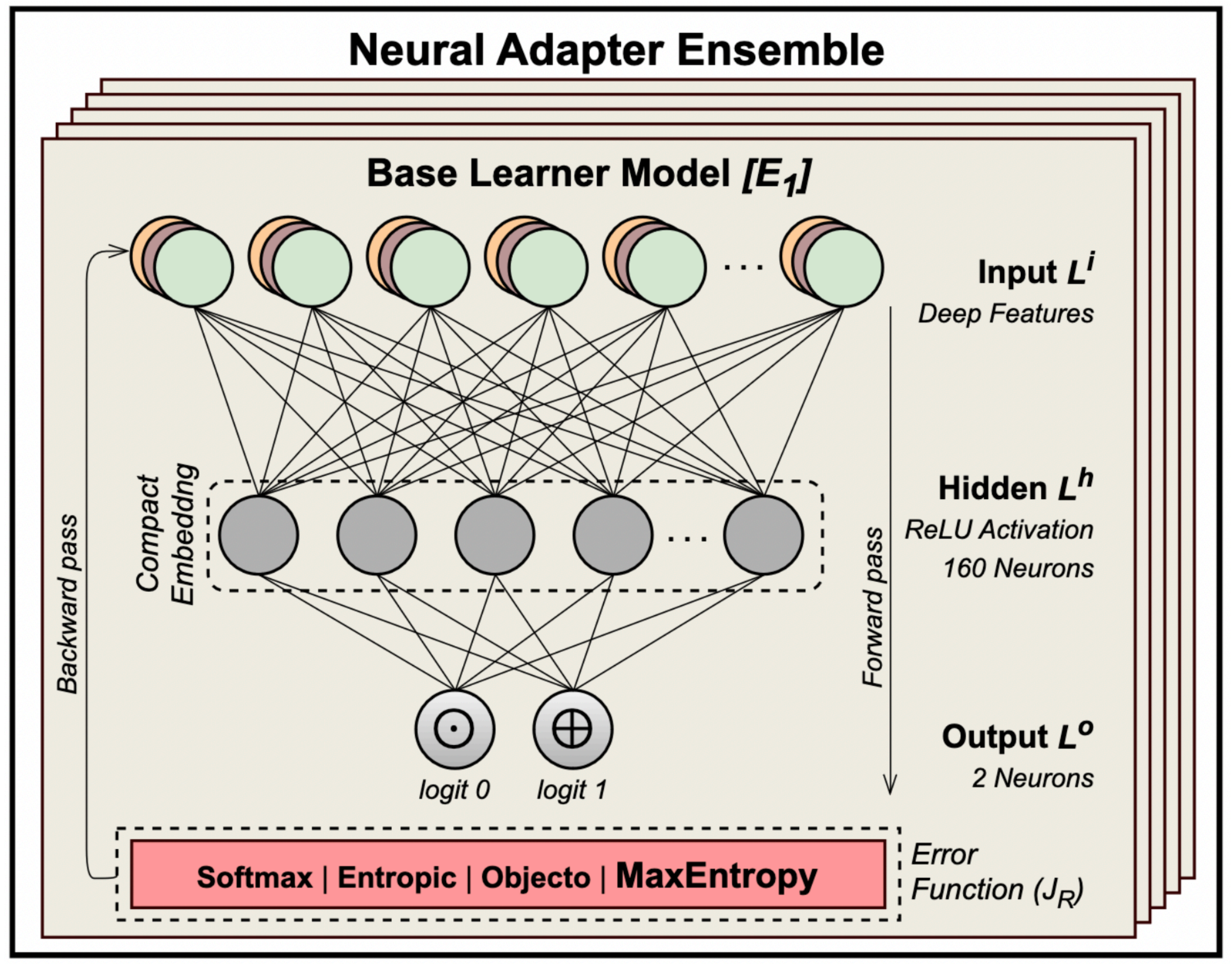




IARPA Janus Benchmark C

EVALUATION: CONTRASTING LOSS FUNCTIONS ON NAE

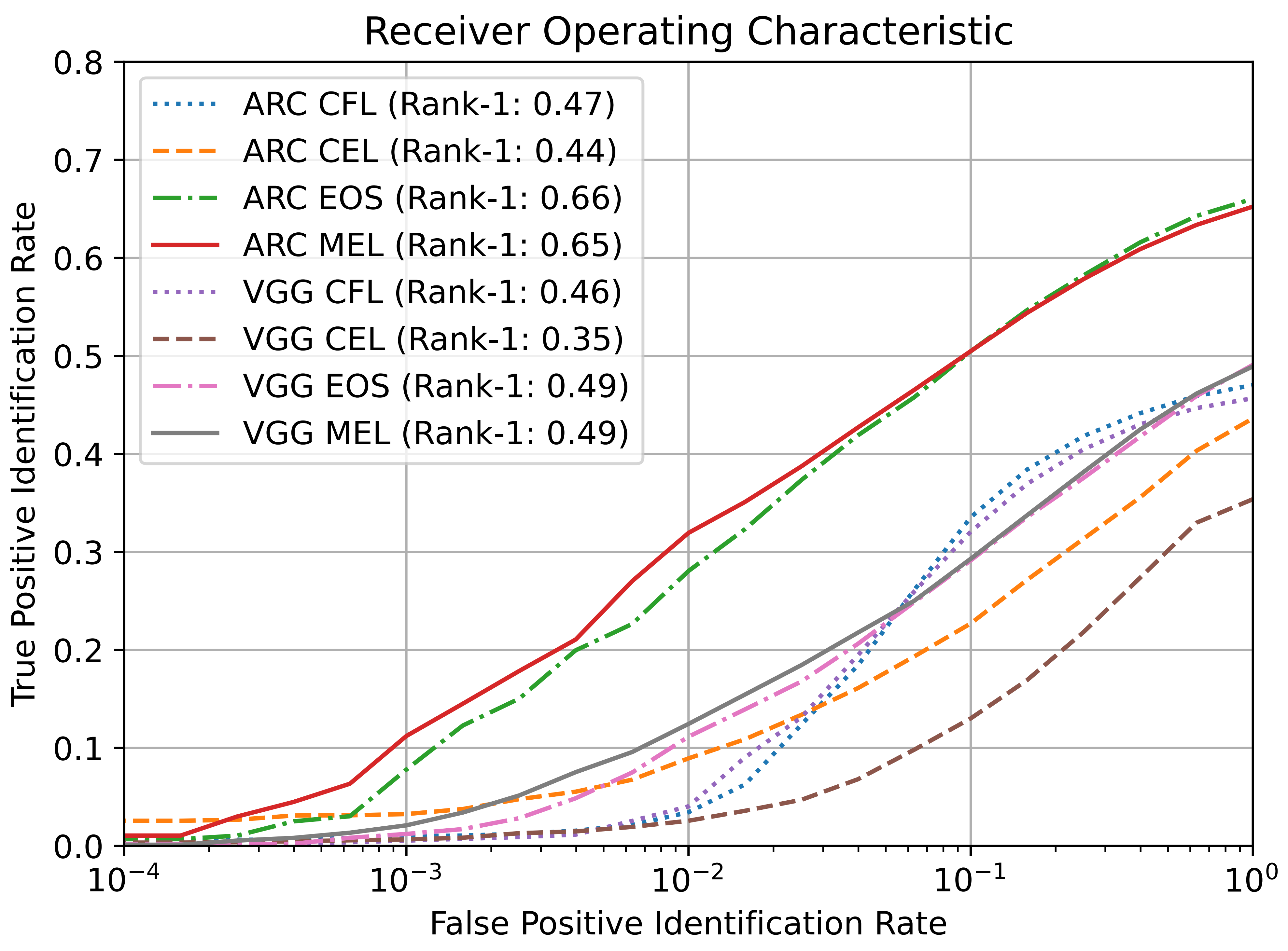




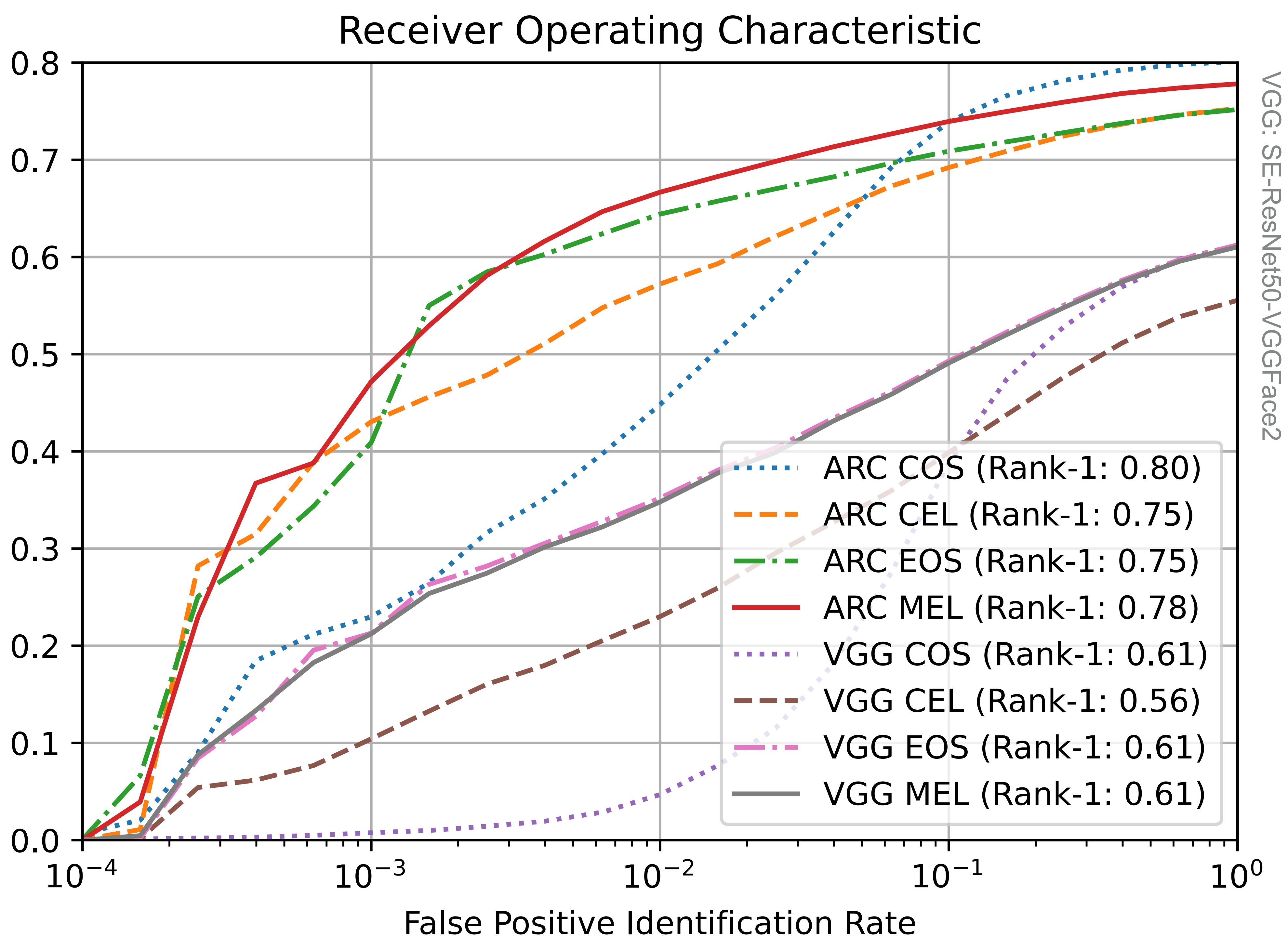
IARPA Janus Benchmark C

EVALUATION: CONTRASTING NEURAL ADAPTER APPROACHES

Neural Adapter Network (NAN)



Neural Adapter Ensemble (NAE)



ARC: ResNet100-ArcFace
VGG: SE-ResNet50-VGGFace2

CONCLUDING REMARKS

- Three different approaches:
  - Cost Function, Feature Augmentation and Ensemble
  - Parameters obtained in one domain and tested in another
  - Negative synthetic data boost performance when properly exploited by cost functions
- What from now?
  - Vision Transformers + MEL + OMU

IF YOU HAVE ANY PERSISTENT QUESTION, REACH US!

HTTP://SMARTSENSELAB.DCC.UFMG.BR/EN/

RAFAEL H. VARETO

RAFAELVARETO@DCC.UFMG.BR