One-Shot Representational Learning for Joint Biometric and Device Authentication

Sudipta Banerjee and Arun Ross

25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
Milan, Italy 10 | 15 January 2021

iprobe.cse.msu.edu
@iPRoBeLab

Paper ID: 1640

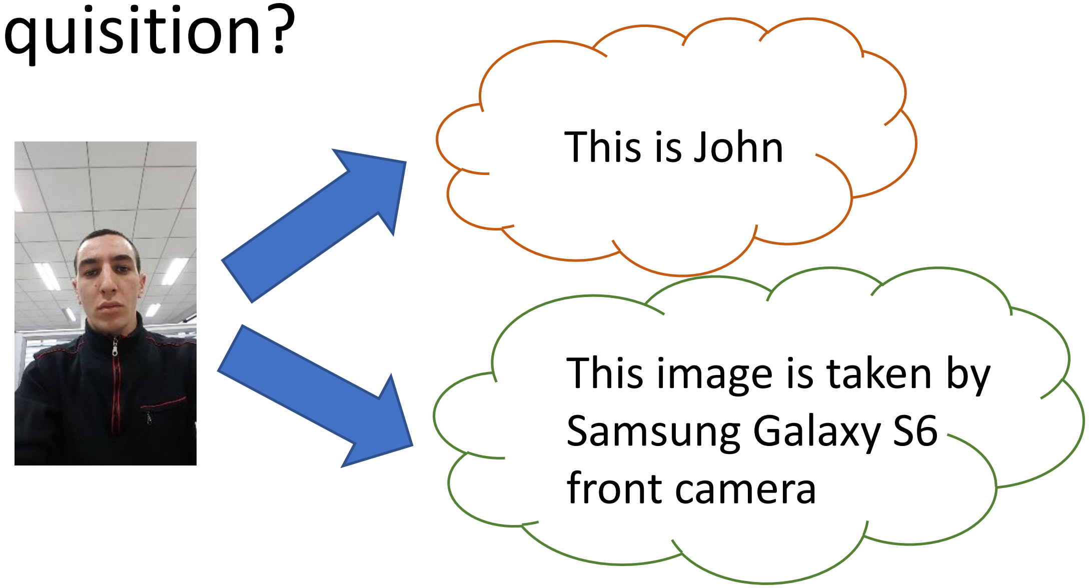
Outline

- Problem Statement
- Motivation
- Proposed Method
- Evaluation
- Summary and Future Work

Problem Statement

Given a biometric image, can we simultaneously identify

- Individual possessing the biometric attribute?
- Device used for acquisition?



Joint biometric-sensor recognition

Motivation

- Smartphone authentication using biometric traits have become common
  - Samsung Galaxy Note 8 uses iris scanner, Apple iPhone X uses Face ID
- Two-factor authentication ensures authorized user and registered device access remote banking
  - By 2023, there will be roughly 2.6 B biometric payment users[1]

[1] How Biometric Authentication is Shaping the Future of Security in Mobile Banking

Image Sources: https://www.samsung.com/in/support/mobile-devices/what-is-the-basic-principle-of-iris-scanning-in-galaxy-note8/ ; https://support.apple.com/en-us/HT208109

Existing Methods

- Media access control (MAC) addresses are unique networking addresses – Can be broadcasted (MAC Flooding, ARP Spoofing)
- SRAM cells as physically unclonable functions – Hardware-based solution
- Separate biometric and device identification modules – Limited by weakest performing module

Joint representation is required that can generalize across different biometric modalities and sensors

Proposed Method

- We explored joint space (high frequency) which captures both biometric-specific and sensor-specific details

$\approx \mathcal{E}(\boldsymbol{X}, k)$ Embedding Network

$$J(\boldsymbol{X}) = B(\boldsymbol{X}) \cap S(\boldsymbol{X})$$

Joint representation

Biometric representation

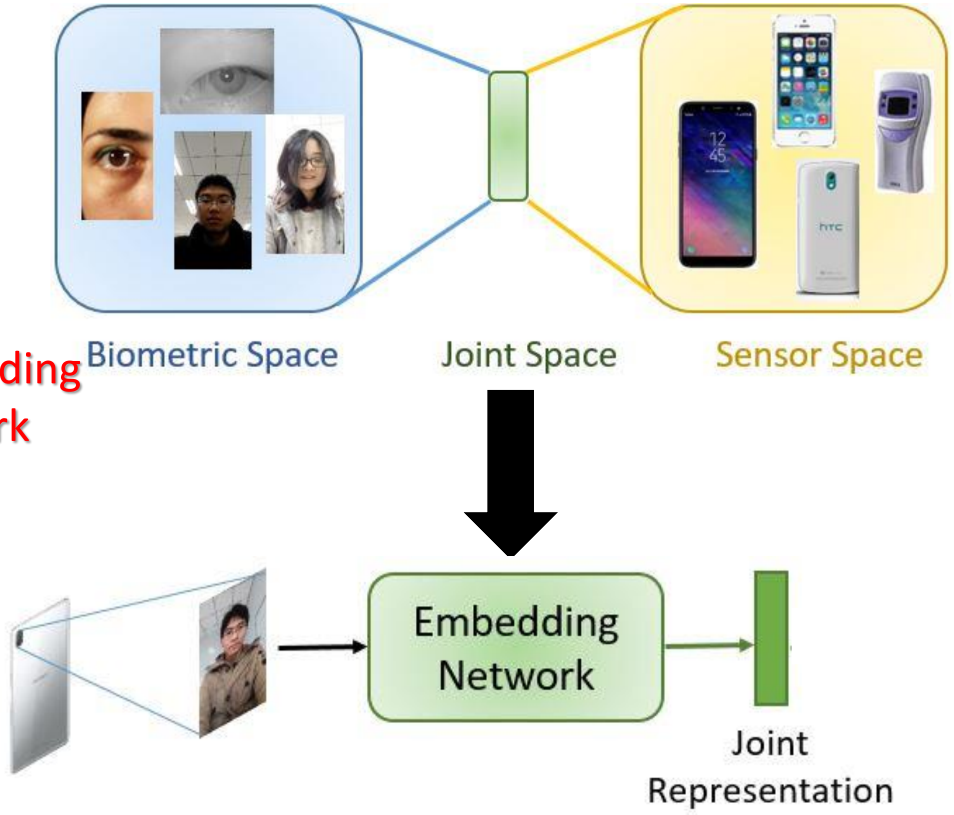
Sensor representation

Input biometric image ➡ $\boldsymbol{X} \in \mathbb{R}^d$

$B(\boldsymbol{X}) \in \mathbb{R}^m$

$S(\boldsymbol{X}) \in \mathbb{R}^n$ $\quad k \leq (m+n) \text{ and } k \ll d$

$J(\boldsymbol{X}) \in \mathbb{R}^k$

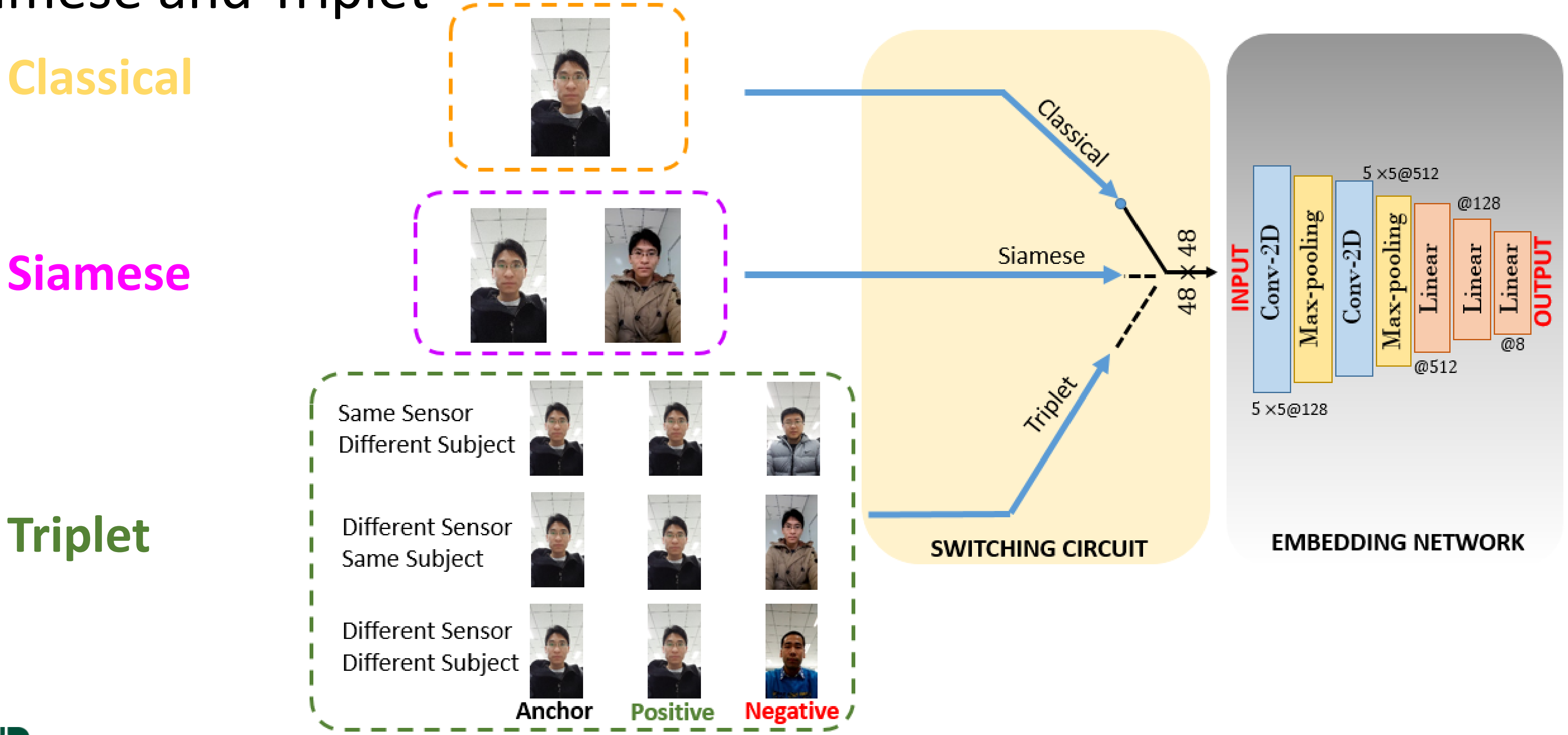


PROPOSED METHOD

- Joint representation – implicitly privacy preserving and embedded in lower dimensional space

Implementation

- Used embedding networks trained in three paradigms – Classical, Siamese and Triplet



Evaluation

- Used three biometric modalities – Iris (near-infrared), Periocular (visible) and Face (visible)

Modality	Dataset	Name of sensors	(# Subjects, # Sensors, # Classes)	Split	# Images
Iris	CASIA-Iris V2	CASIA IrisCAM-V2, OKI IrisPass-h	(60, 2, 120)	Train Test	1,680 720
Periocular	MICHE-I	Apple iPhone 5S (Front and Rear) UNIT I and UNIT II, Samsung Galaxy S4 (Front and Rear), Samsung Galxy Tab GT2 (Front)	(75, 7, 375*)	Train Test	2,278 863
Face	OULU-NPU	HTC Desire EYE, Sony XPERIA C5 Ultra Dual, MEIZU X5, Oppo N3, Samsung Galaxy S6 Edge, ASUS Zenfne Selfie	(55, 6, 330)	Train Test	5,940 2,970
TOTAL			(190, 13, 825)		14,451

- Employed different loss functions for handling multiple negative classes
- Conducted evaluation in joint identification and joint verification scenarios

Training mode	Loss function	
Classical	Cross entropy	
Siamese	Single margin contrastive loss (SMCL) Double margin contrastive loss (DMCL)	
Triplet	Offline triplet mining	
	Online triplet mining	Random negative Semi-hard negative Hardest negative
	Multi-class N-pair	All positive pair Hard negative pair

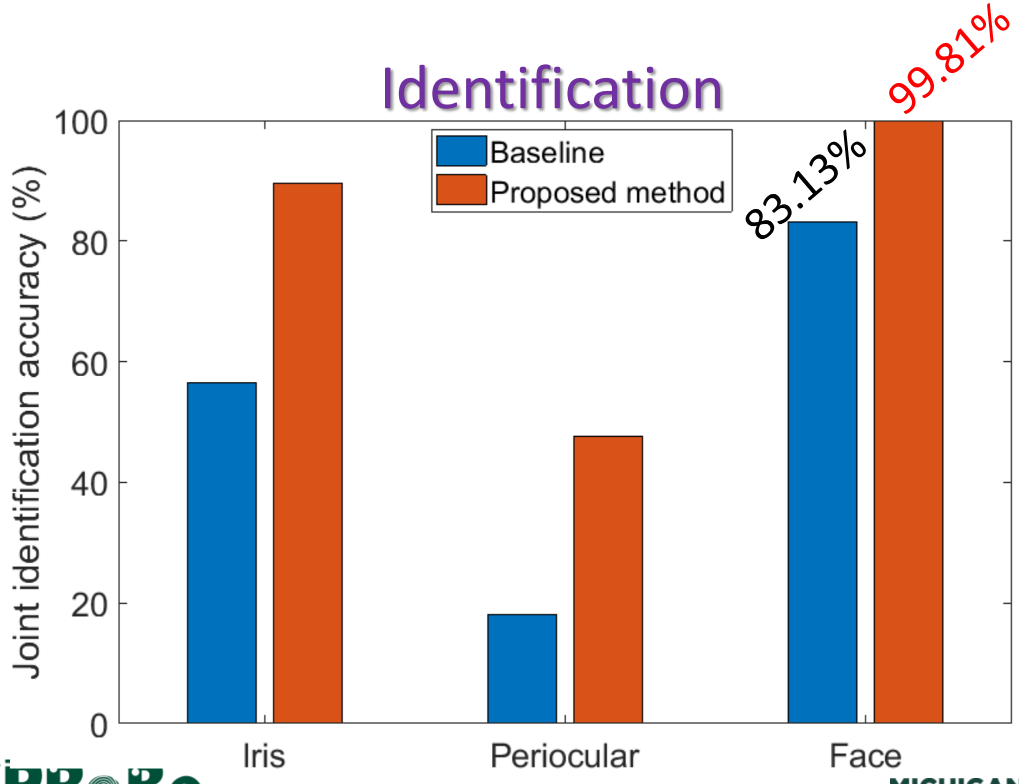
Evaluation (contd.)

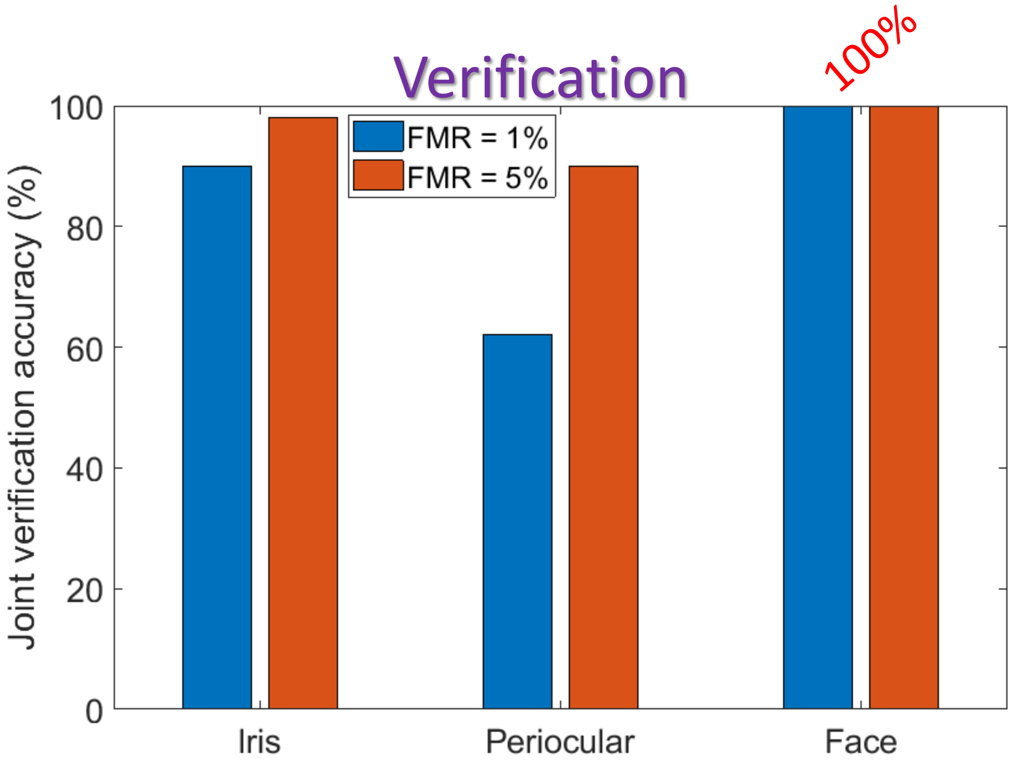
- Joint identification/verification – Only if both subject and device identities yield correct matches
- Baselines – Only for joint identification
  - Enhanced Photo Response Non-Uniformity[1] – Sensor (Device) identification
  - COTS Iris and Face matchers, ResNet 101 Off-the-shelf Periocular matcher – Biometric identification
- Evaluations conducted for different
  - image resolution,
  - joint representation dimensionality
  - distance metrics for embedding
- Best results obtained for 48x48 input image, 8-dimensional embedding and standardized Euclidean distance as metric

[1] C. T. Li, "Source camera identification using enhanced sensor pattern noise," *T-IFS*, 2010

Results and Analysis

- Training module – Siamese embedding network with Single-Margin Contrastive loss outperformed all other combinations of networks and loss functions used in this work





iPRoBe

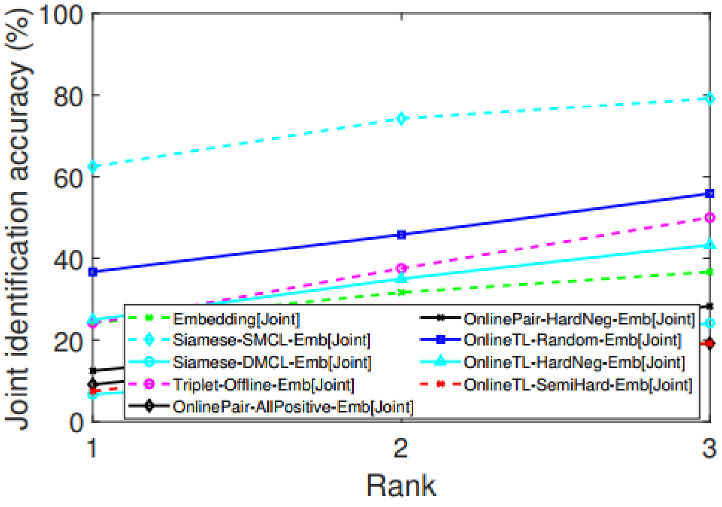
MICHIGAN STATE UNIVERSITY

Chopra et al., “Learning a similarity metric discriminatively, with application to face verification ,” CVPR 2005

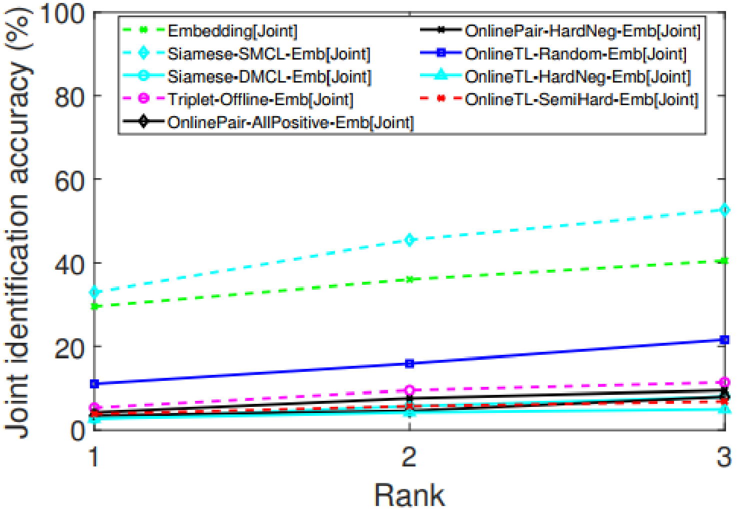
Paper ID: 1640

Results and Analysis (contd.)

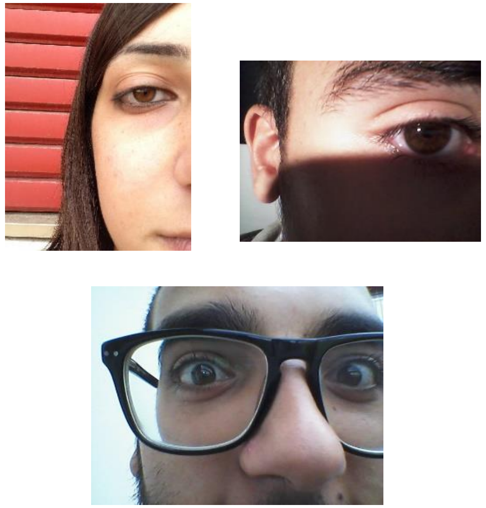
- Investigated causes for relatively worse performance on periocular (MICHE-I) images in the task of joint identification
  - Performance variation across left and right eyes
  - Challenges – illumination variations (indoor and outdoor) and occlusions



LEFT EYE



RIGHT EYE



CHALLENGING IMAGES

Marsico et al., "Insights into the results of MICHE I - mobile iris challenge evaluation," PR 2018

Summary and Future Work

- Developed a joint representation that can be used to simultaneously authenticate the user and the device (sensor)
- Results indicate performance up to 99.81% (@ Rank 1) in terms of joint identification and up to 100% (@ False Match Rate 1%) in terms of joint verification
- Method generalizes across Iris, Face and Periocular biometric modalities and sensors operating in visible and near-infrared spectra
- Future work will focus on large-scale evaluation and performance improvement for periocular images

Thank You!
Q&A

Contact: banerj24@msu.edu