HDIAC

Homeland Defense & Security
Information Analysis Center

Night-time Face Recognition

Shuowen (Sean) Hu, Ph.D.
Nathan Short, Ph.D.
Ben Riggan, Ph.D.

U.S. Army Research Laboratory

September 20, 2018

The views presented are those of the speaker and do not necessarily represent the views of DoD or its components.

Introduction

HDIAC & Today's Topic

HDIAC Overview

What is the Homeland Defense & Security Information Analysis Center (HDIAC)?

One of three Department of Defense Information Analysis Centers

Responsible for acquiring, analyzing, and disseminating relevant scientific and technical information, in each of its eight focus areas, in support of the DoD and U.S. government R&D activities

HDIAC's Mission

Our mission is to be the go-to R&D/S&T and RDT&E leader within the homeland defense and security (HDS) community, by providing timely and relevant information, superior technical solutions, and quality products to the DoD and HDS Communities of Interest/Communities of Practice.

Homeland Defense & Security Information Analysis Center

HDIAC Overview

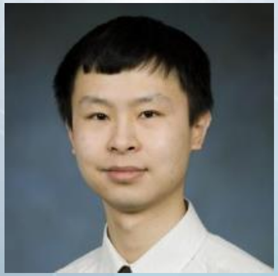
HDIAC Subject Matter Expert (SME) Network

HDIAC SMEs are experts in their field(s), and, typically, have been published in technical journals and publications.

SMEs are involved in a variety of HDIAC activities
- Authoring HDIAC Journal articles
- Answering HDIAC Technical Inquiries
- Engaging in active discussions in the HDIAC community
- Assisting with HDIAC Core Analysis Tasks
- Presenting webinars

If you are interested in applying to become a SME, please visit HDIAC.org or email info@hdiac.org.

Presenters



Shuowen (Sean) Hu, Ph.D.
Dr. Shuowen (Sean) Hu received the B.S. in electrical and computer engineering from Cornell University in 2005, and a Ph.D. in electrical and computer engineering from Purdue University in 2009. He was awarded the Andrews Fellowship to study at Purdue University. Following graduation from Purdue University, he joined the U.S. Army Research Laboratory (ARL) as an electronics engineer in the Image Processing Branch. His current research focus is on cross-spectrum face recognition as well as on target detection and classification.

Nathan Short, Ph.D.
Dr. Nathan Short is a Lead Scientist at Booz Allen Hamilton. He received his M.S. and Ph. D. degrees in Computer Engineering from Virginia Tech in 2012. He conducts research and development in computer vision, image processing and machine learning, supporting government organizations within DoD, DHS, DOJ, and the IC. His experience includes R&D of imaging systems for unmanned vehicles, multi-biometric solutions for mobile and traditional assets as well as developing next generation human identification technology to support forensic and ISR applications.

Ben Riggan, Ph.D.
Dr. Ben Riggan received the B.S. degree in computer engineering from N.C. State University in 2009, and M.S. and Ph.D. degrees in electrical engineering from N.C. State University in 2011 and 2014, respectively. After finishing his Ph.D., he was awarded a postdoctoral fellowship at the U.S. Army Research Laboratory’s Image Processing Branch, where he worked on face recognition. Currently, he works for the Networked Sensing and Fusion Branch of the U.S. Army Research Laboratory. Dr. Riggan's research interests are in areas of biometrics and fusion, which leverage his expertise in image/signal processing, computer vision, and machine learning.

Overview

Biometrics

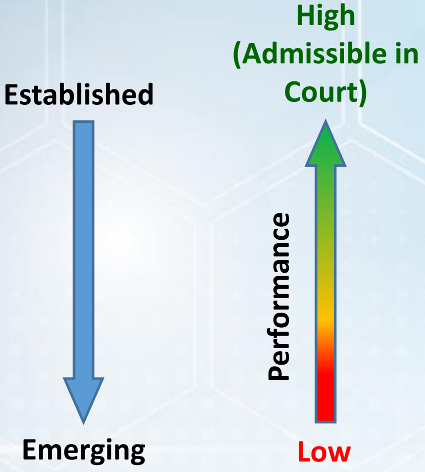
Merriam-Webster dictionary: Measurement and analysis of unique physical behavioral characteristics, especially as a means of verifying personal identity.

Seven factors (Jain et al., 1997):

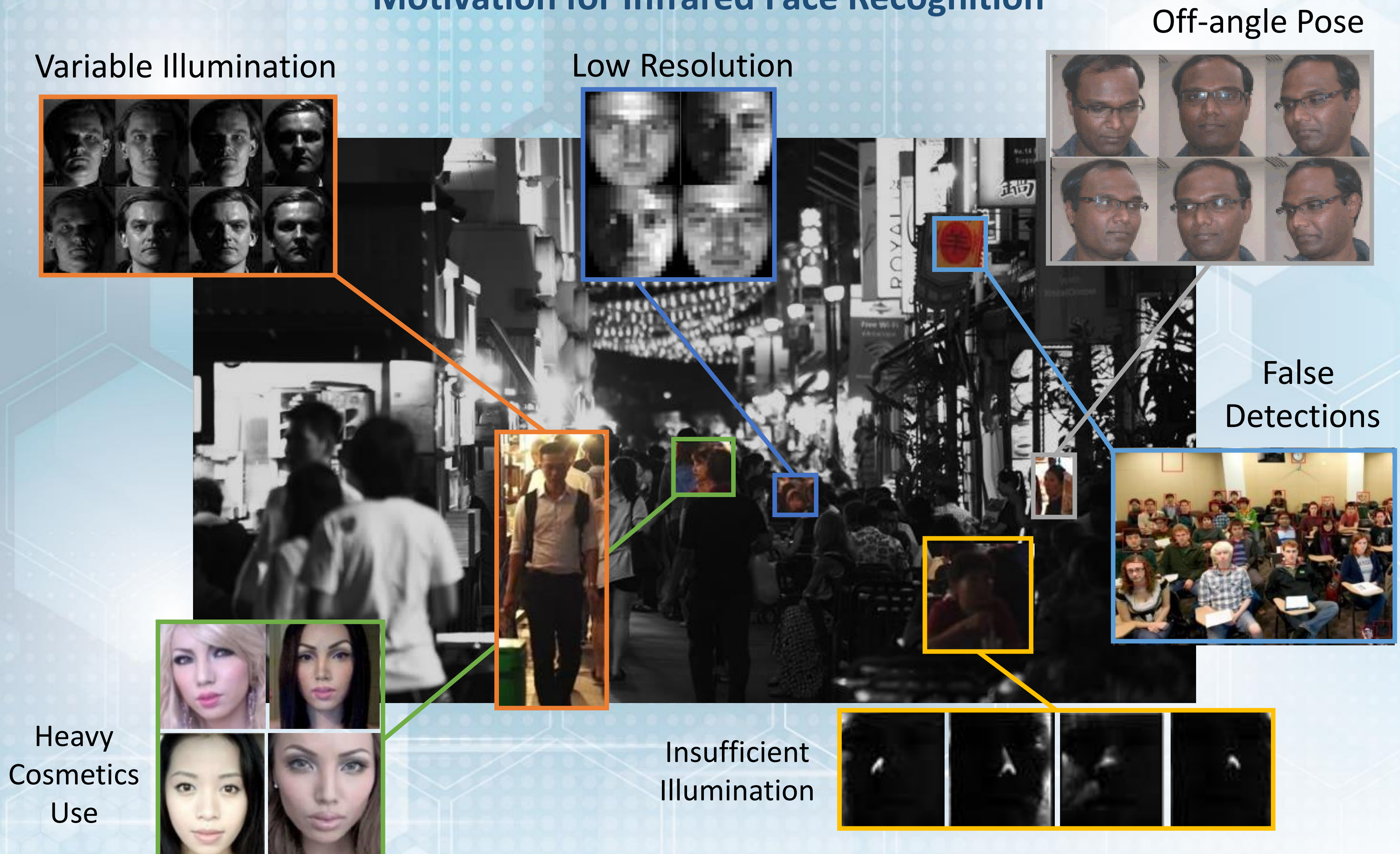
- Universality: every individual should possess the trait
- Uniqueness: discriminative between individuals
- Permanence: degree of invariance to the passage of time
- Measurability/collectability
- Performance
- Acceptability
- Circumvention (e.g., spoofing, presentation attack)

Biometric modalities:

- DNA
- Fingerprint
- Iris
- **Face**
- Voice
- Soft biometrics: gait/anthropometry, scars marks tattoos (SMT), ear



Motivation for Infrared Face Recognition



Homeland Defense & Security Information Analysis Center

Infrared Face Recognition

Exploit infrared facial signatures for cross-spectrum face recognition, matching against visible spectrum gallery databases

Benefits:

- Infrared (IR) light invisible to human eye, illumination independent
- Near (N) and Short Wave (SW) IR features highly correlated with visible band, but require active illumination (detectable) in low-light environments
- Mid Wave (MW) and Long Wave (LW) IR sensors operate passively without external illuminator, lack corresponding details with visible
- Polarimetric-thermal provides more detail through passive acquisition

VIS	NIR	SWIR	MWIR	LWIR	Polarimetric
0.4-0.75μm	0.75-1.4 μm	1.4-3 μm	3-8 μm	8-15 μm	7.5-11.1μm

Recognition Performance

Well-Correlated

Acquisition Covertness

Fully Passive

Polarimetric Face Recognition

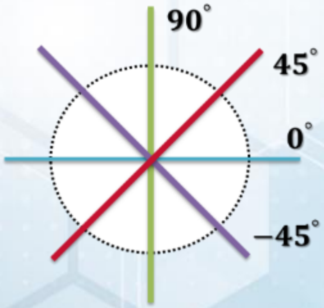
Advantages:

Polarimetric LWIR provides key textural and geometric facial details not present in conventional thermal face signature

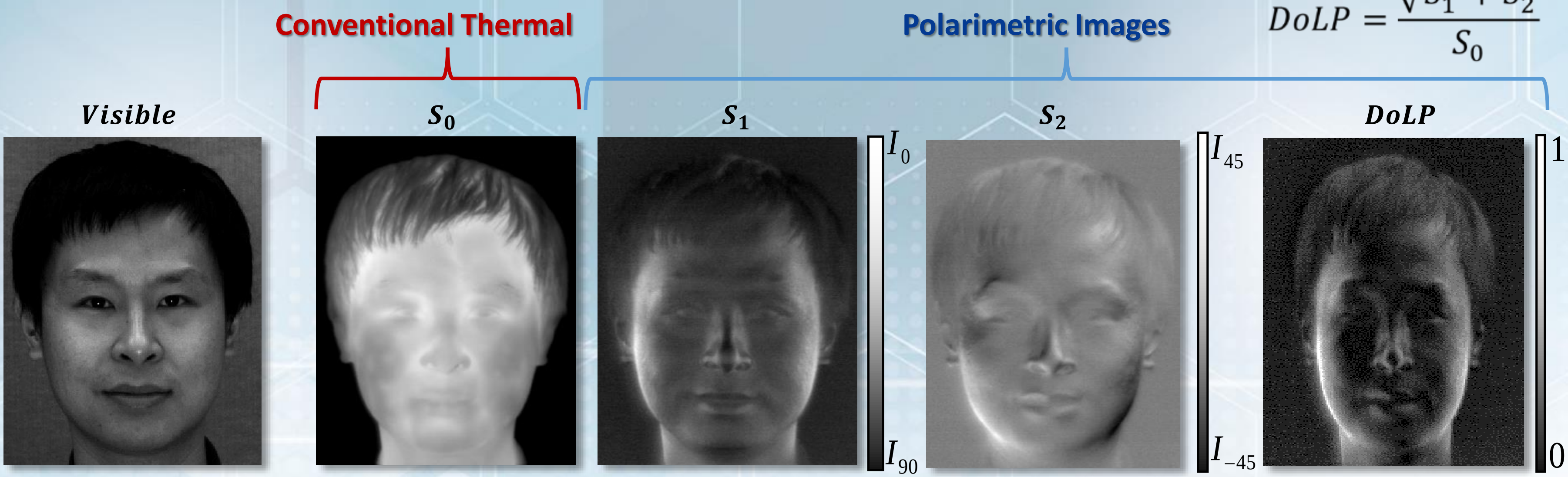
Polarimetric characteristics:

- Measures emission intensity at different polarization-states
- Stokes vectors describe preferred polarization-state of captured light
- Degree of Linear Polarization (DoLP) used to approximate amount of linearly polarized light emitting from a source
- Provides information about surface texture and orientation of surface normal with respect to viewing angle

Stokes Vector



$$S_0 = I_0 + I_{90}$$
$$S_1 = I_0 - I_{90}$$
$$S_2 = I_{45} - I_{-45}$$
$$DoLP = \frac{\sqrt{S_1^2 + S_2^2}}{S_0}$$



Invariance to Makeup/Cosmetics

The problem of cosmetic changes in appearance:

- Facial cosmetics significantly compromises face recognition algorithm performance in the visible spectrum (Eckert et al., 2013; Chen et al., 2013; Dantcheva et al., 2012)
- Polarimetric thermal imaging is dependent on surface normal orientation, and is therefore relatively unaffected by application of make-up, as initial experiments show

CAMOUFLAGE PAINT — **Without and with camouflage paint**

COSMETIC MAKEUP — **Without and with cosmetic makeup**

EXTREME DISGUISE — **Without and with extreme disguise**

Visible Spectrum

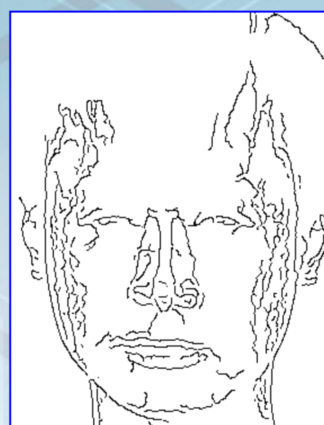
Polarimetric

Edge Maps

Canny(vis)

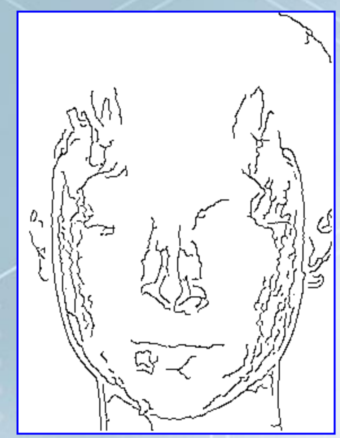
Canny(S_0)

Canny(S_1)



Canny(S_2)

Canny(*DoLP*)

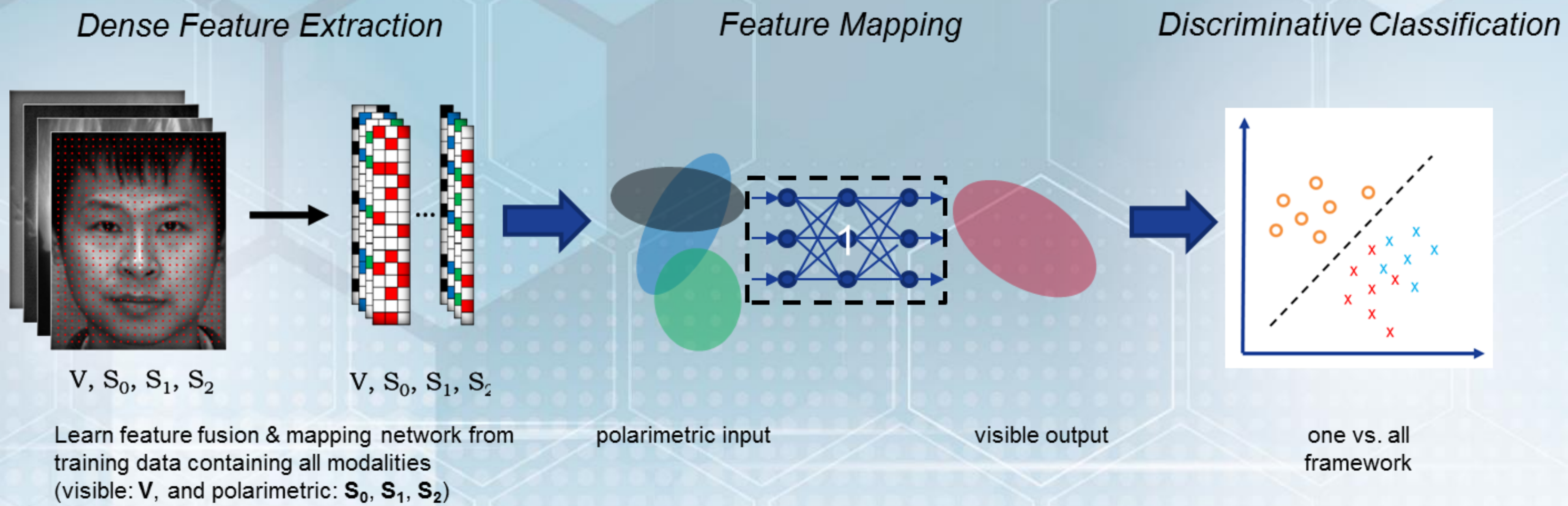


Key edges are correlated between thermal and visible facial signatures; polarization state information complements intensity-only thermal information

Learning Cross-spectrum Features

Advanced feature mapping and fusion improves cross-modal polarimetric thermal-to-visible face recognition

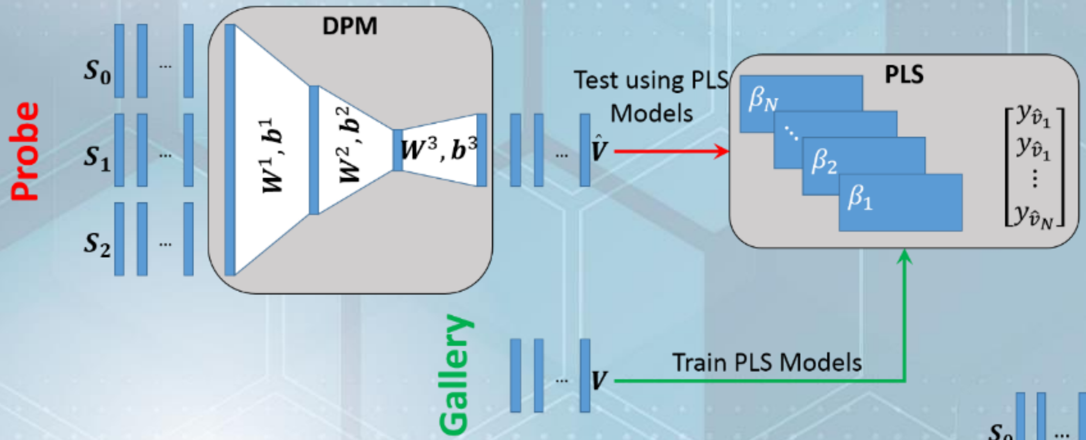
- Perform direct mapping of polarimetric thermal features to visible feature space for enhanced recognition
- Further enhance performance by adding a discriminative classifier to recognition framework



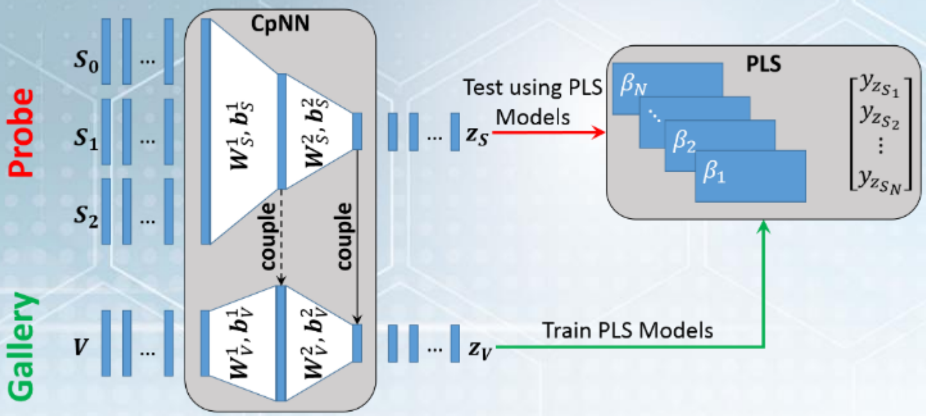
Feature Learning and Discriminative Framework

Approaches:

- Learn mapping from Polarimetric feature space to Visible feature space using shallow neural networks:
    - Deep Perceptual Mapping (DPM) (**direct regression)**



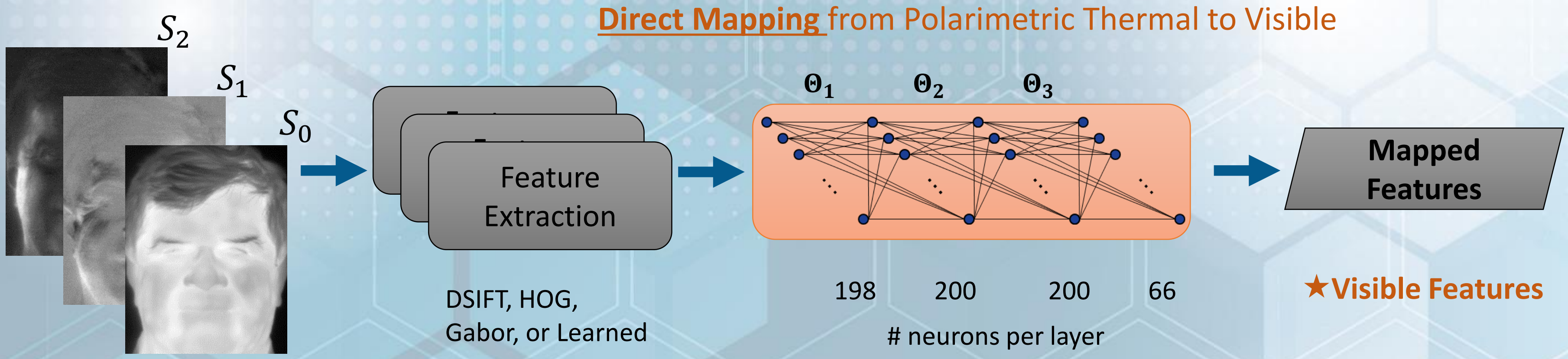
- Coupled Neural Network (CpNN) (**indirect regression)**



Homeland Defense & Security Information Analysis Center

Deep Perceptual Mapping

Direct Mapping from Polarimetric Thermal to Visible



Data Preparation:

- Landmark detection, alignment, and cropping
- Extract features from Polarimetric Thermal and Visible (**training only!**) images

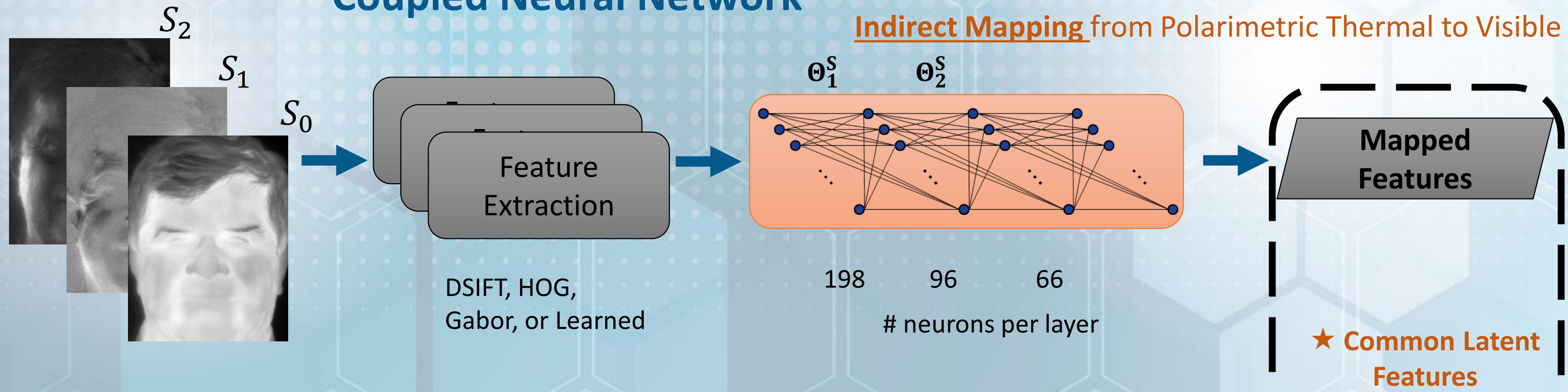
Optimization:

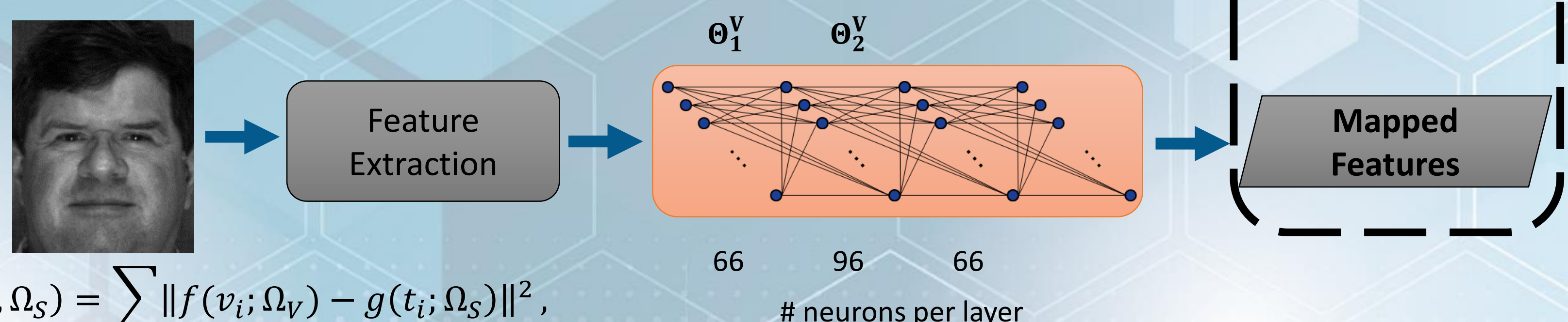
$$J(\Omega) = \sum_i \|v_i - g(t_i;\Omega)\|^2 ,$$

where

v_i	—	visible
t_i	—	polarimetric
$g(\cdot)$	—	learned mapping
Ω	—	model params.

Coupled Neural Network





Optimization:

$$J(\Omega_V, \Omega_S) = \sum_i \|f(v_i; \Omega_V) - g(t_i; \Omega_S)\|^2 ,$$

where

v_i	—	visible
t_i	—	polarimetric
$f(\cdot)$	—	visible mapping
$g(\cdot)$	—	polarimetric mapping
Ω_V	—	visible model params.
Ω_S	—	polarimetric model params.

- ★ Gallery – mapped visible (BOTTOM ONLY)
- ★ Probes – mapped polarimetric (TOP ONLY)
- ★ Trained using preprocessed/corresponding visible and polarimetric thermal imagery.

Face Verification Performance

Feature based mapping and matching techniques for cross-spectrum face recognition

- Improves thermal-to-visible face recognition with conventional FLIR imagers in the thermal spectrum
- Improves polarimetric thermal-to-visible face recognition by exploiting polarization state information acquired using maturing polarimetric imaging technology

Verification Performance
(Range 1 baseline)

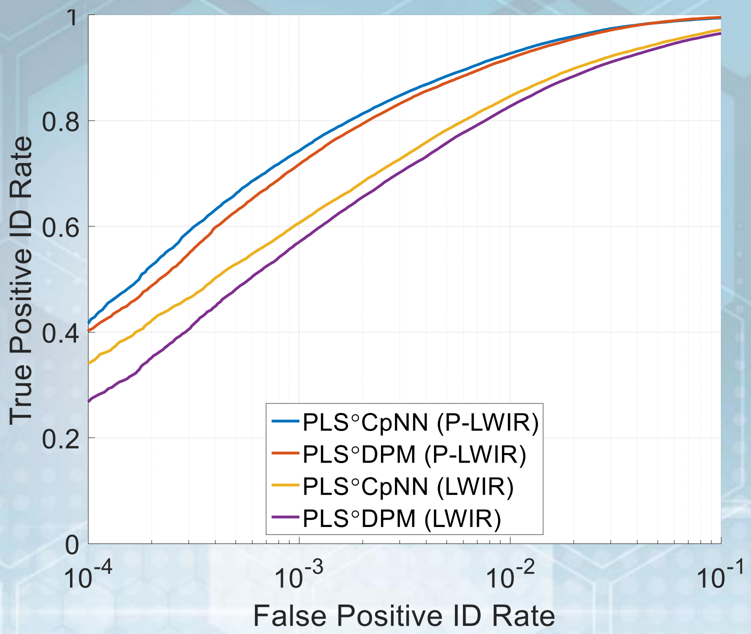
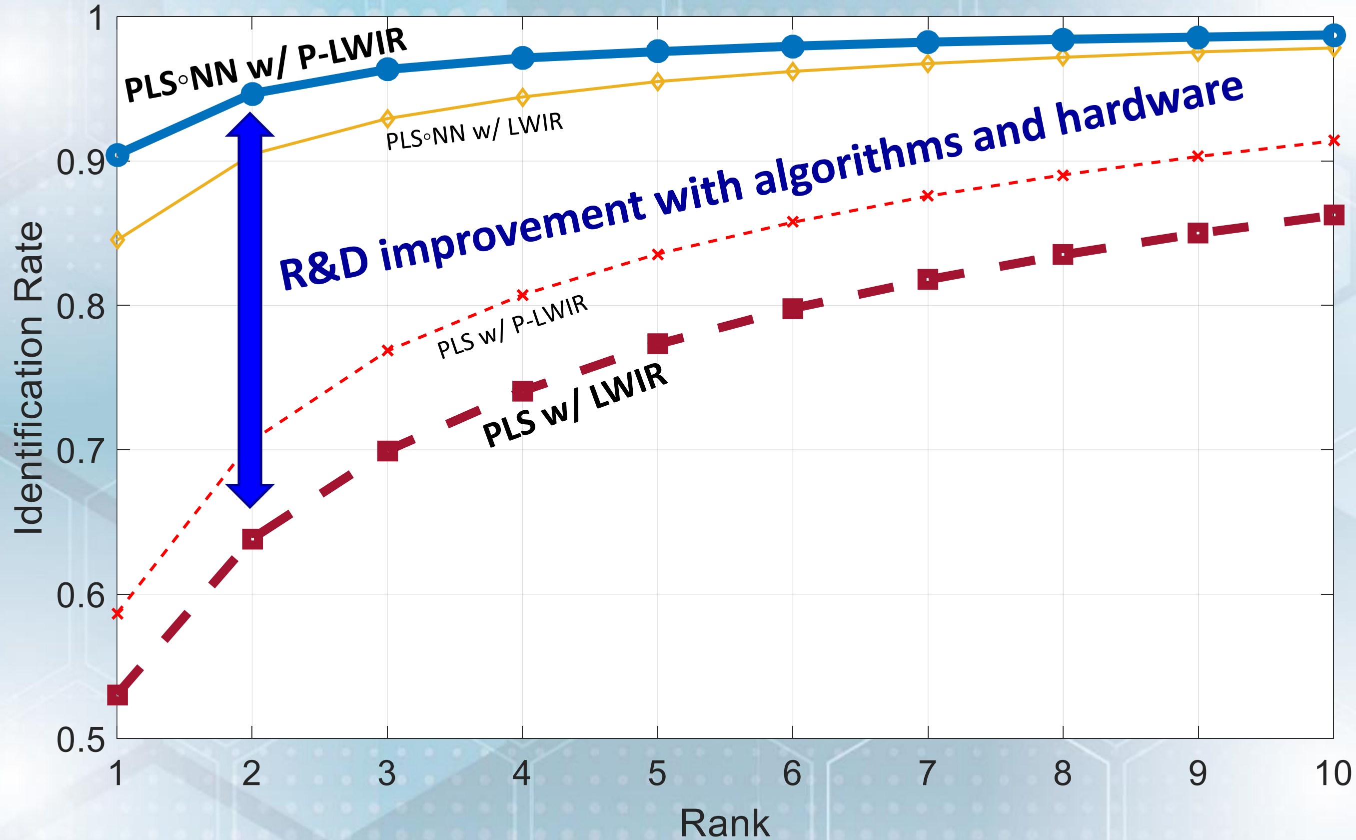


Figure and table show current results using two feature mapping techniques (PLS◦CpNN & PLS◦DPM) for:

- Conventional thermal-to-visible recognition (labeled LWIR, achieves 84.8% true positive ID rate at 1% false positive ID rate)
- Polarimetric thermal-to-visible recognition (labeled P-LWIR, achieves 92.8% TPIR at 1% FPIR)

Method	TPIR @FPIR=0.01	TPIR @FPIR=0.001
$PLS \circ CpNN\ (\boldsymbol{P-LWIR})$	0.928	0.743
$PLS \circ DPM\ (\boldsymbol{P-LWIR})$	0.918	0.719
$PLS \circ CpNN\ (LWIR)$	0.848	0.607
$PLS \circ DPM\ (LWIR)$	0.828	0.574

Identification Performance



Results: 25-subject training dataset, 35-subject testing dataset

- Feature mapping followed by discriminative classification improved polarimetric thermal-to-visible Rank-1 ID from **59%** (PLS) to:
  - **90%** with neural network mapping followed by PLS matching (PLS∘NN Map)

Motivation for Cross-Spectrum Synthesis

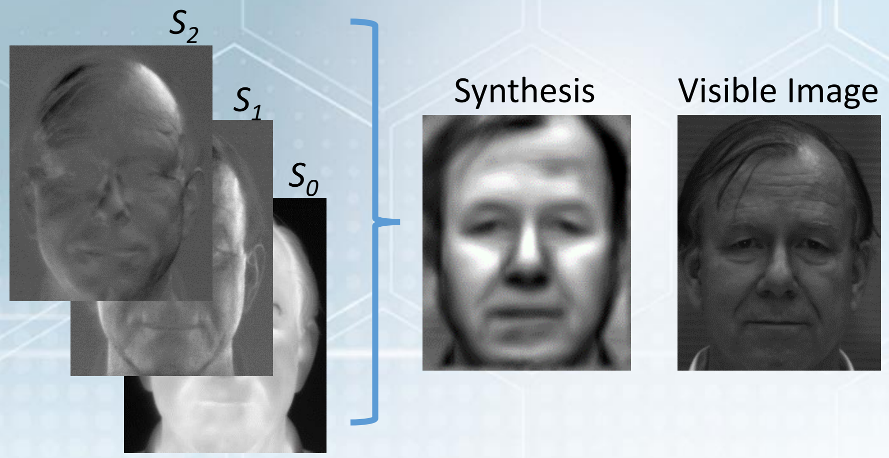
Objective: Synthesis a visible-like face image from a polarimetric thermal *input*, generating photo-realism WHILE preserving discriminative characteristics.

Returned Matches

Automatic Face Recognition
Daytime & Nighttime

Human-in-the-loop
Analysts need to be able to identify correct matches

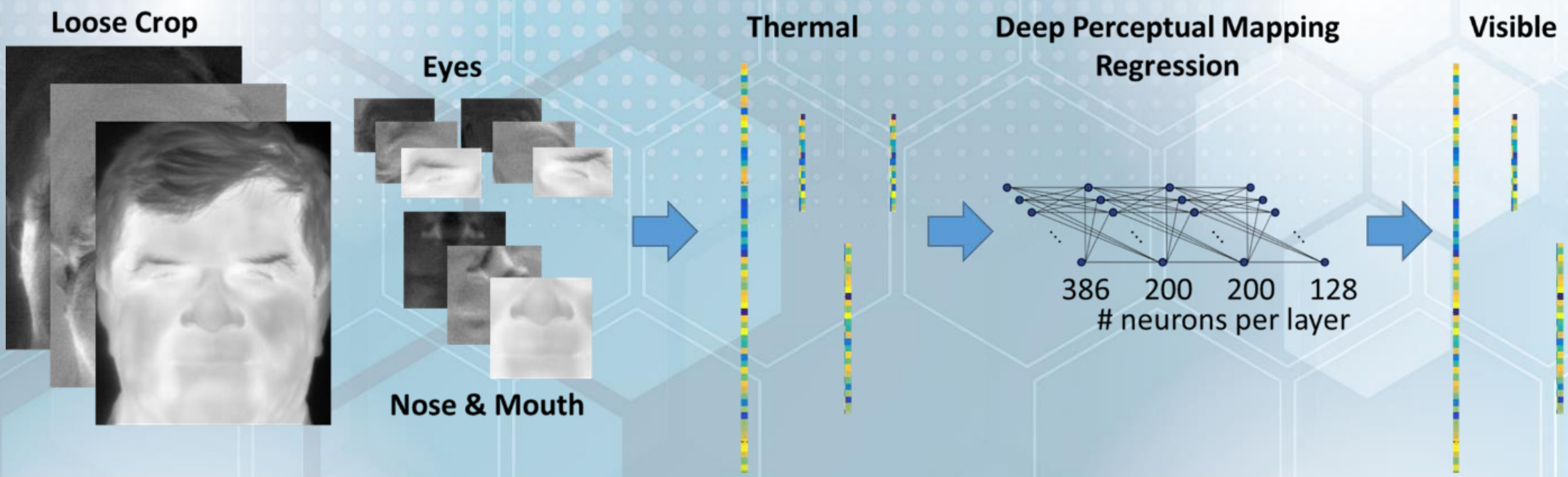
Verification



Advantages of Synthesis:

- **Leverage COTS & GOTS face recognition systems (trained on visible spectrum)**
- **Enables human examiner adjudication**

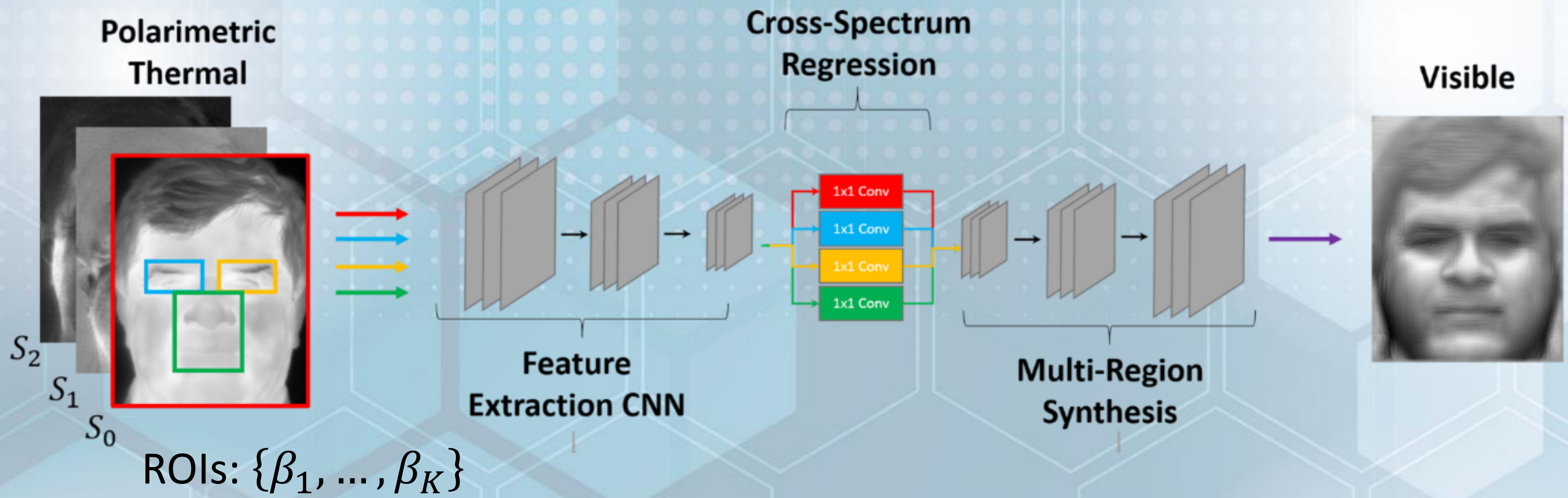
Multi-Region Synthesis



Multi-Region, Cross-Spectrum Regression…

- Extract features from regions: face, eyes, nose, and mouth (4 regions)
- Domain Adaptation — thermal-to-visible Deep Perceptual Mapping (DPM) [Riggan et al. 2016, Sarfraz et al. 2017]
- One DPM per region.

Multi-Region Synthesis



Joint Optimization

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}} J(\boldsymbol{x}) = \arg\min_{\boldsymbol{x}} \sum_{i=1}^{K} \omega_i J_i(\boldsymbol{x})$$

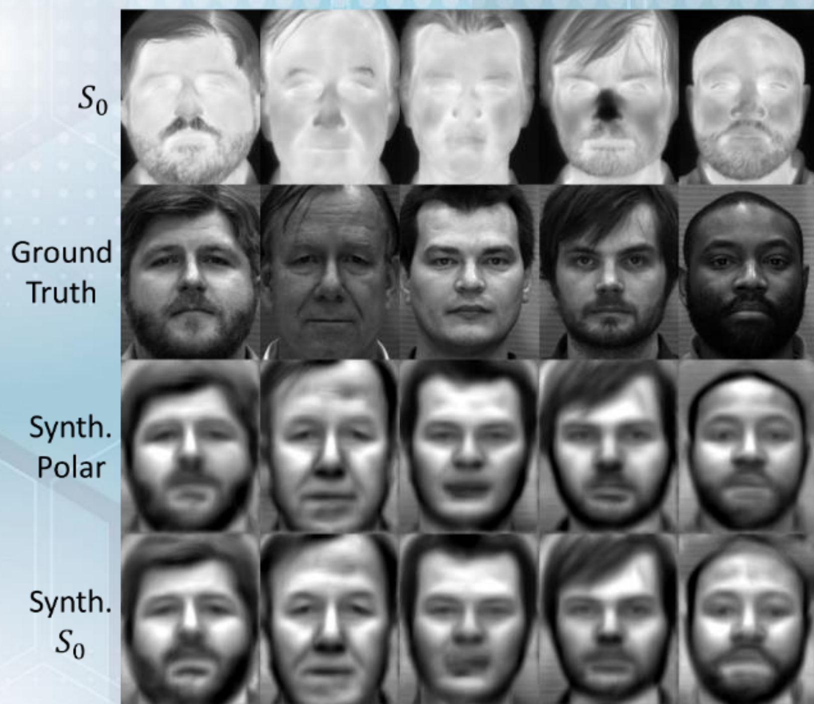
where

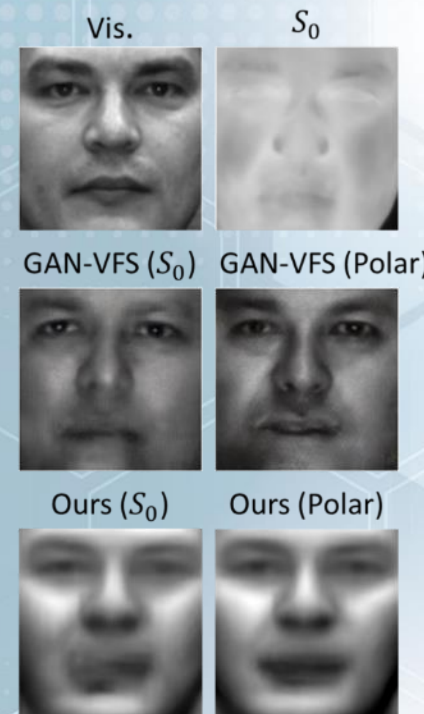
$$J_i(\boldsymbol{x}) = \begin{cases} \|g(\tilde{\boldsymbol{x}}_i) - h_i \circ g(\tilde{\boldsymbol{t}}_i)\|^2 + \lambda R(\tilde{\boldsymbol{x}}_i) & , \tilde{x}_i = \{x_{u,v} : (u,v) \in \beta_i\} \\ 0 & , otherwise \end{cases}$$

(1) alpha norm — $\|\boldsymbol{x}\|_\alpha^\alpha$

(2) total variation — $\sum_{i,j}\left(\left(x_{i,j+1} - x_{i,j}\right)^2 + \left(x_{i+1,j} - x_{i,j}\right)^2\right)^{\beta/2}$

Results



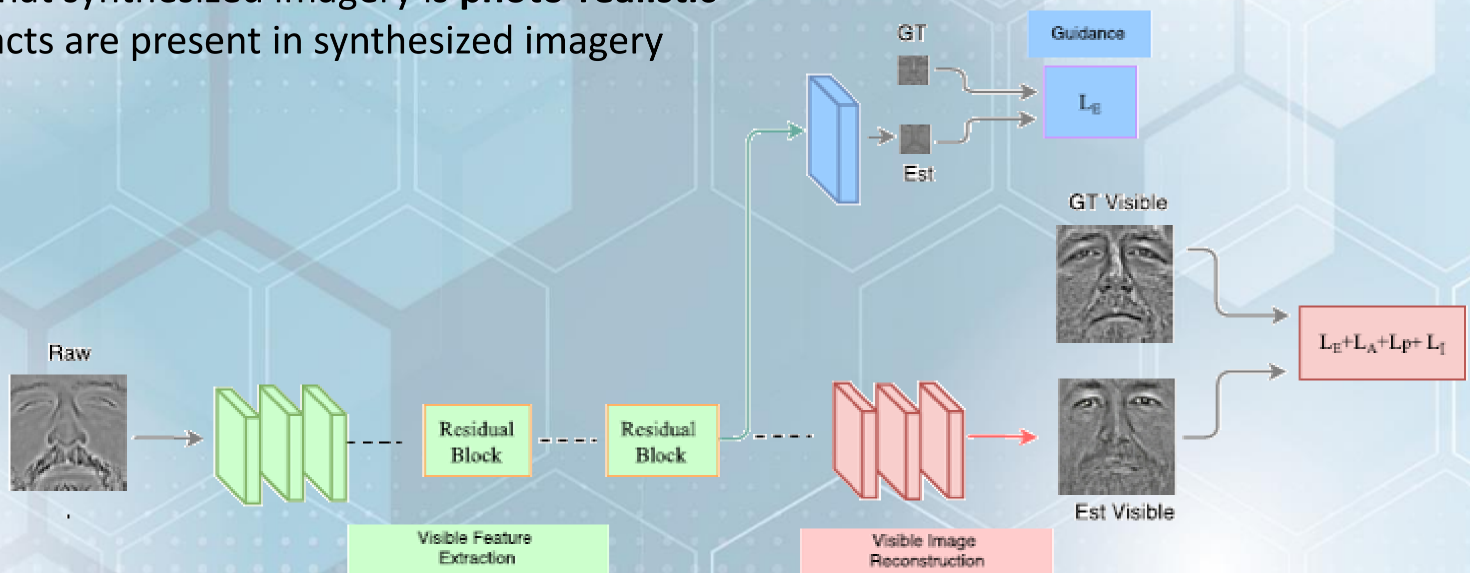


Method	AUC (polar)	AUC (S_0)	EER (polar)	EER (S_0)
Raw	50.35%	58.64%	48.96%	43.96%
Mahendran et al. 2014	58.38%	59.25%	44.56%	43.56%
Riggan et al. 2016	75.83%	68.52%	33.20%	34.36%
Zhang et al. 2017	79.90%	79.30%	25.17%	27.34%
Multi-Region Synthesis (**ours**)	**85.43%**	**82.49%**	**21.46%**	**26.25%**

Landmark Detection using DLIB

GAN-based Synthesis

- Cross-spectrum synthesis approach using Generative Adversarial Networks (first proposed by Ian Goodfellow).
- Collaboration with Rutgers University (Professor Vishal Patel and He Zhang)
- Observed that synthesized imagery is **photo-realistic**
- Some artifacts are present in synthesized imagery



Loss Function

- L_2 loss on guidance sub-network
- L_2 loss on recovered visible image
- Adversarial loss: $-\log(\phi_D(\phi_G(\cdot))$
- Perceptual loss: L_2 loss on relu3-1 layer (pretrained VGG model)
- Identity loss: L_2 loss on relu2-2 layer (fine-tuned VGG-Polar model)
- $L_{total} = L_{L_2} + L_{L_2(G)} + \lambda_A L_A + \lambda_P L_P + \lambda_I L_I$

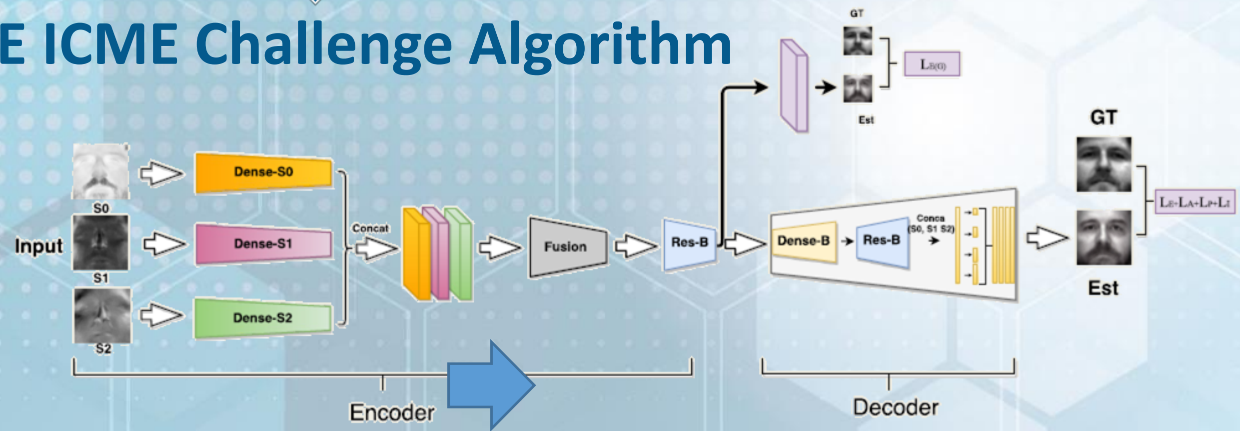
GAN-based Synthesis Results

Ground Truth

GAN Results (multiple losses)

IEEE ICME Challenge Algorithm

Polarimetric Thermal Input



Generator

Synthesized

Ground Truth Visible

Discriminator

"Generative Adversarial Network Based Multi-Stream Fusion for Cross-Spectrum Synthesis & Recognition"

- **He Zhang and Vishal Patel**

IEEE ICME Challenge Algorithm

Team: **He Zhang (Rutgers University) & Professor Vishal Patel (now with Johns Hopkins)**

GAN based multi-stream feature-level fusion for Synthesis:

- Generator is a multi-stream encoder-decoder network using dense-residual blocks
- A deep guided subnet is stacked at the end of the encoder, incorporating perceptual loss and identity preserving loss in addition to adversarial loss
- Multi-scale patch-discriminator

Matching:

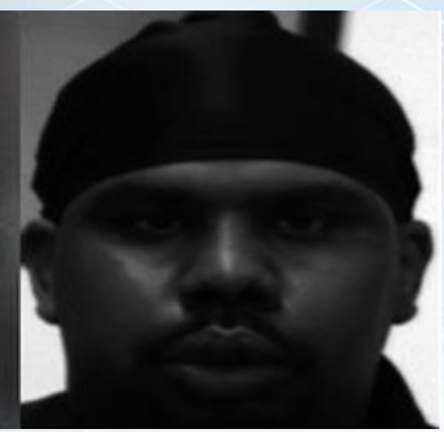
- VGG-Face used to extract features from visible face image and synthesized face image (from polarimetric thermal input)
- Cosine similarity used to produce match score

Synthesized

Visible

Synthesized

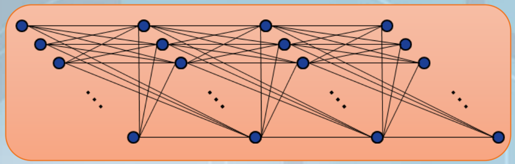
Visible

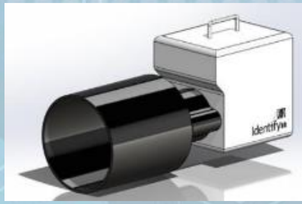


From Theory to Practice

Thrust 1: HFR Algorithm Development

- Develop neural networks based HFR algorithms to exploit polarimetric thermal signature and match against visible gallery
- Develop IR face and fiducial point (e.g., eyes, nose) detection techniques
- **Engage with academia and industry**



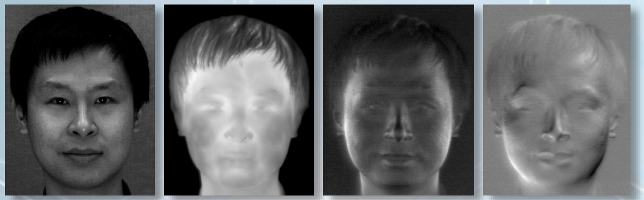


Thrust 4: Sensors

- **Leverage advances in thermal imagers** (i.e., COTS systems). Decreasing cost.
- Through SBIR program, develop next generation of polarimeters
  - New design to mitigate motion artifacts and improve sensitivity
  - Custom optics for standoff acquisition

Thrust 2: Demonstrator System

- **Understand user requirements, obtain stakeholder feedback**
- Develop demonstration prototype for polarimetric thermal-to-visible FR
  - Streamed polarimetric thermal data
  - Real-time face & fiducial point detection
  - Real-time cross-spectrum matching



Thrust 3: Data collection

- Collect multi-spectrum facial signatures to facilitate algorithm development and improve robustness
- **Increased training samples improve feature learning**
- Multi-spectrum facial signatures exhibiting real-world variability improves robustness of learned features

Conclusion

- Challenges for nighttime face recognition:
  - Recognition accuracy using extended gallery: need extensive training data to more effectively leverage recent advances in machine learning
  - Automated and accurate detection of facial points: need to develop fiducial point and face detection algorithms
  - Pose invariance: incorporate frontalization and off-pose training data
  - Extended standoff range
- Path forward:
  - **Government, industry, and academia S&T collaboration**
    - Data collection
    - Algorithm development
    - Sensor advancement
  - **Engage operational community**

Questions?