2017 IEEE International Conference on Image Processing

SSPP-DAN: Deep Domain Adaptation Network for Face Recognition with Single Sample Per Person

Sungeun Hong, Woobin Im, Jongbin Ryu, Hyun S. Yang

AIM Lab, KAIST, South Korea

Contents

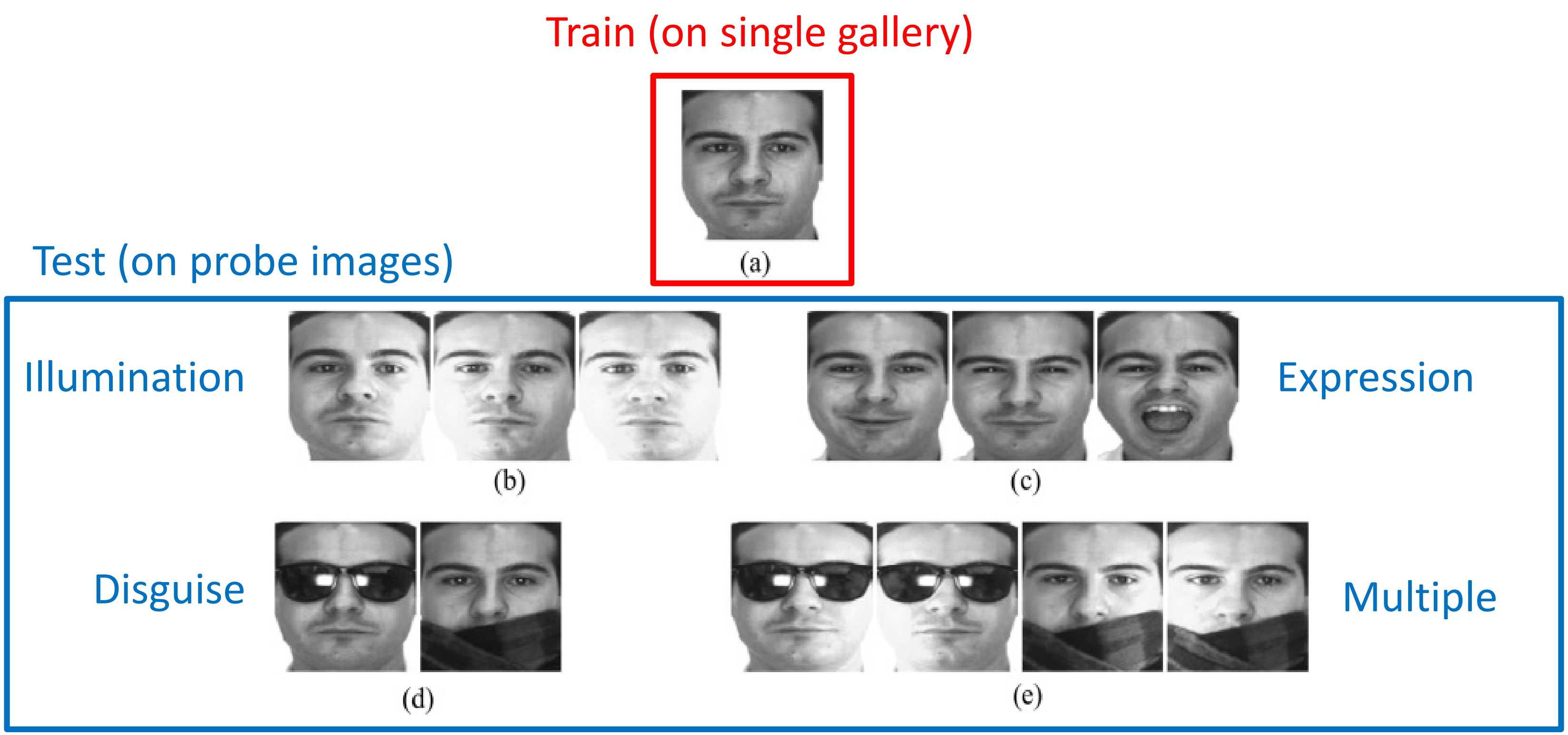
- **Motivation & Problem Definition**
  - Single sample per person (SSPP)
  - Challenges in real-world face recognition
- **Proposed method**
  - Domain adaptation
  - Face synthesis
- **Experiments**
  - New heterogeneous dataset
  - LFW for SSPP

KAIST

Motivation & Problem Definition

SSPP face recognition

- **Face recognition using Single Sample Per Person (SSPP)**
  - Identify or verify identities using <u>**only one single gallery image**</u>
  - Related to the recently attracted one-shot learning

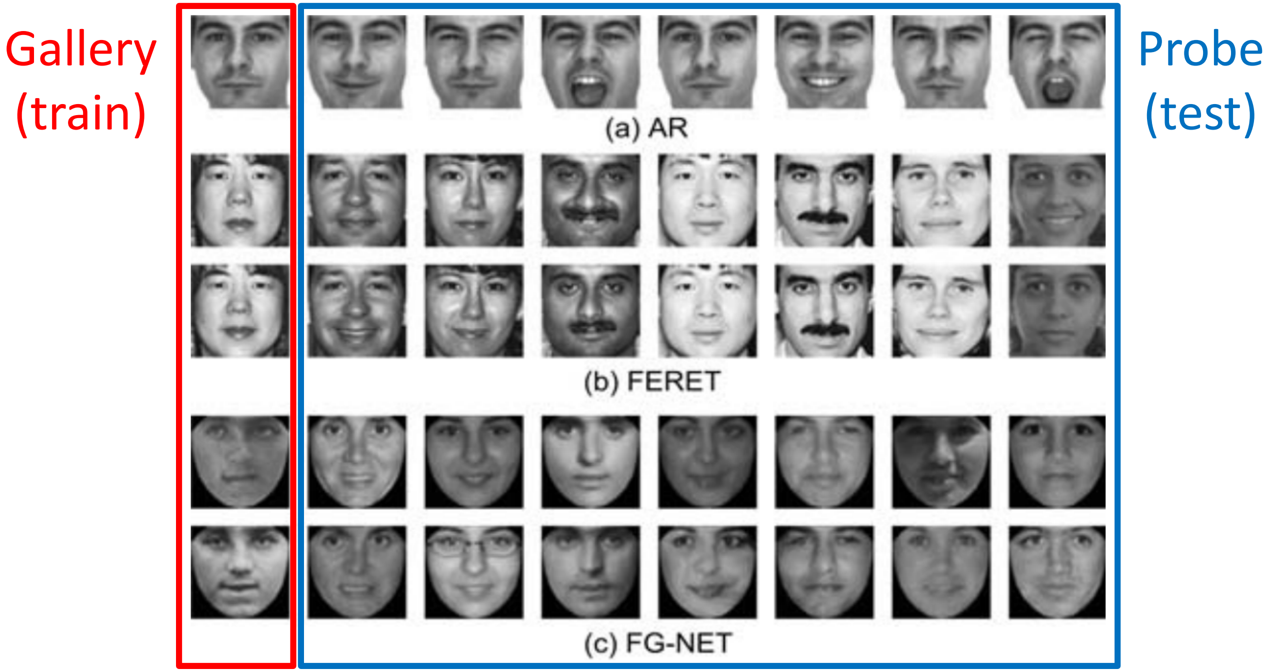


Sample images of the AR database

SSPP face recognition

▪ Limitations of existing SSPP datasets

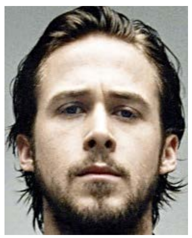
- Lab controlled environment
- Consistent shooting environment



Lu, Jiwen, Yap-Peng Tan, and Gang Wang. "Discriminative multimanifold analysis for face recognition from a single training sample per person." *IEEE transactions on pattern analysis and machine intelligence* 35.1 (2013): 39-51.

Real-world SSPP Face recognition

Registration
(**Gallery**)



Identification
(**Probe**)

Gallery

A stable image like
clear frontal mugshot
e.g., ID card or e-passport

Probe

Unstable images including
non-trivial variations
e.g., surveillance camera, web images

Variations:
camera sensor, blur,
noise, pose, illumination

Real-world SSPP Face recognition

Registration
(**Gallery**)

Identification
(**Probe**)

Challenges

1. Heterogeneity of the shooting environments

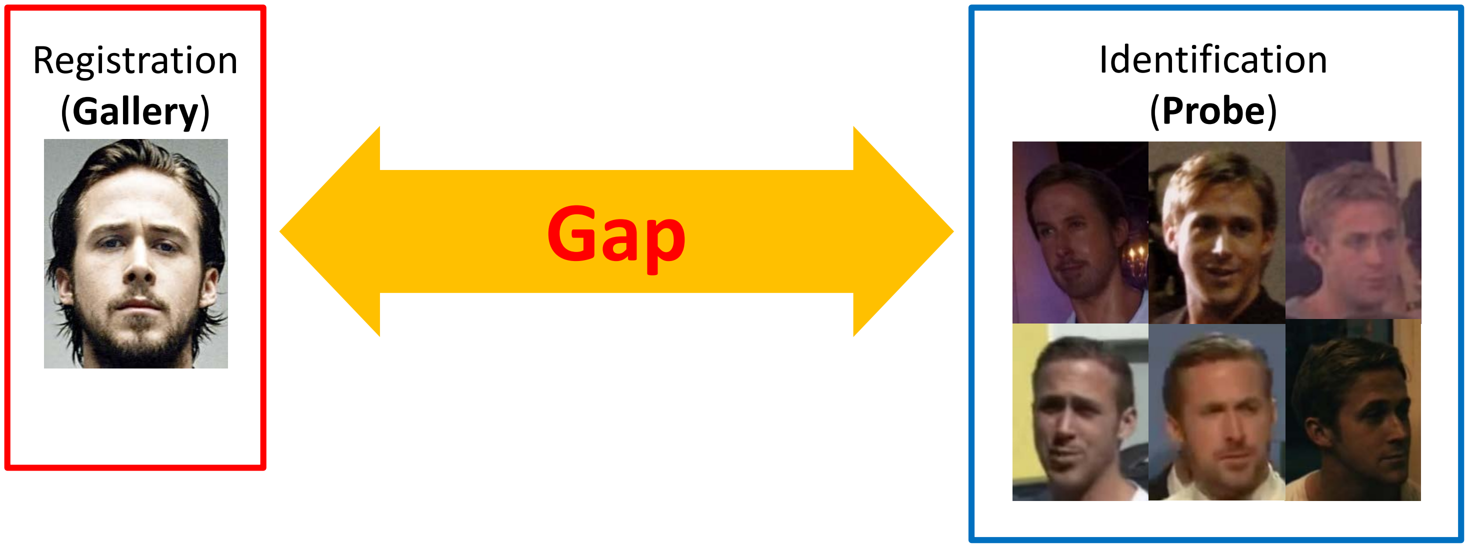
- Gallery: stable environment
- Probe: highly unstable environment

2. Shortage of training samples

- Only one training sample per person is available

Proposed method

Real-world SSPP Face recognition

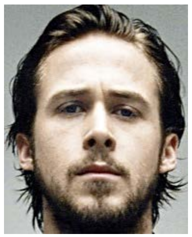


Challenges

1. **Heterogeneity of the shooting environments**
   - Gallery: stable environment
   - Probe: highly unstable environment

Real-world SSPP Face recognition

Registration
(**Gallery**)



Domain Adaptation

Identification
(**Probe**)

Challenges

1. Heterogeneity of the shooting environments

- Gallery: stable environment
- Probe: highly unstable environment

Domain Adaptation (DA)

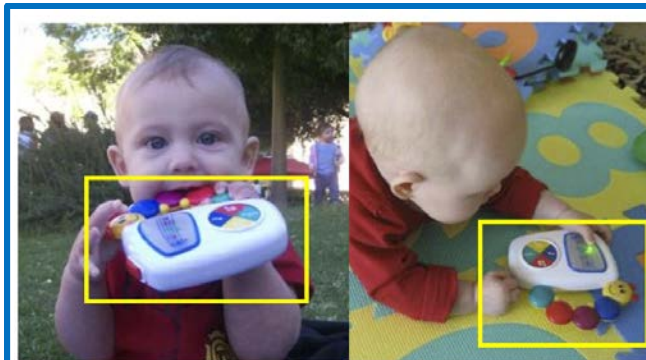
- Adjust a model to a different target domain distribution starting from the source domain knowledge

Source domain

amazon.com

With labels

Target domain



consumer images

Without labels

Domain Adaptation (DA)

- Adjust a model to a different target domain distribution starting from the source domain knowledge

Source domain w/ label

Domain Adaptation (DA)

- Adjust a model to a different <u>target domain distribution</u> starting from the <u>source domain knowledge</u>

Source domain w/ label

Model

Domain Adaptation (DA)

- Adjust a model to a different <u>target domain distribution</u> starting from the <u>source domain knowledge</u>

Source domain w/ label

↓

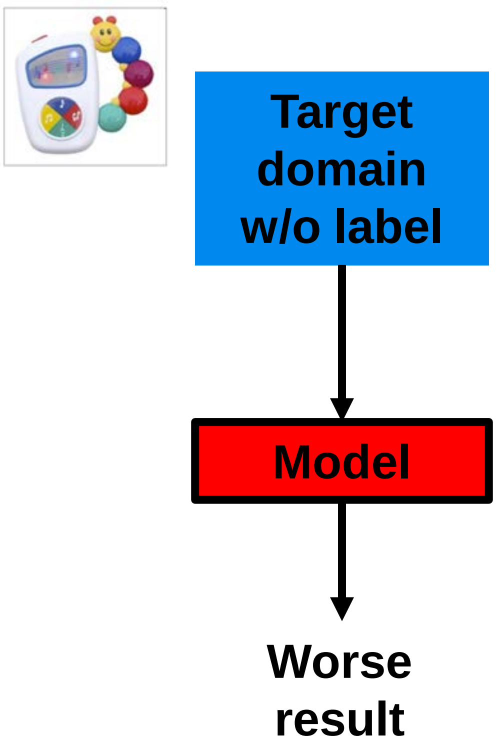
Model

↓

Fine classification

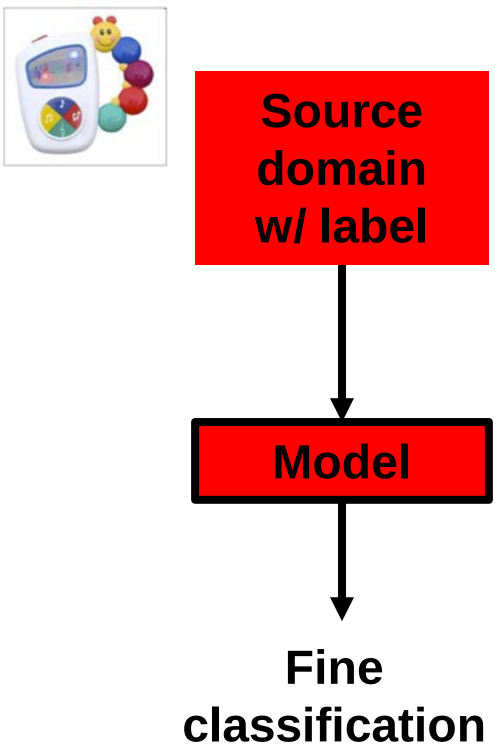
Domain Adaptation (DA)

- Adjust a model to a different target domain distribution starting from the source domain knowledge



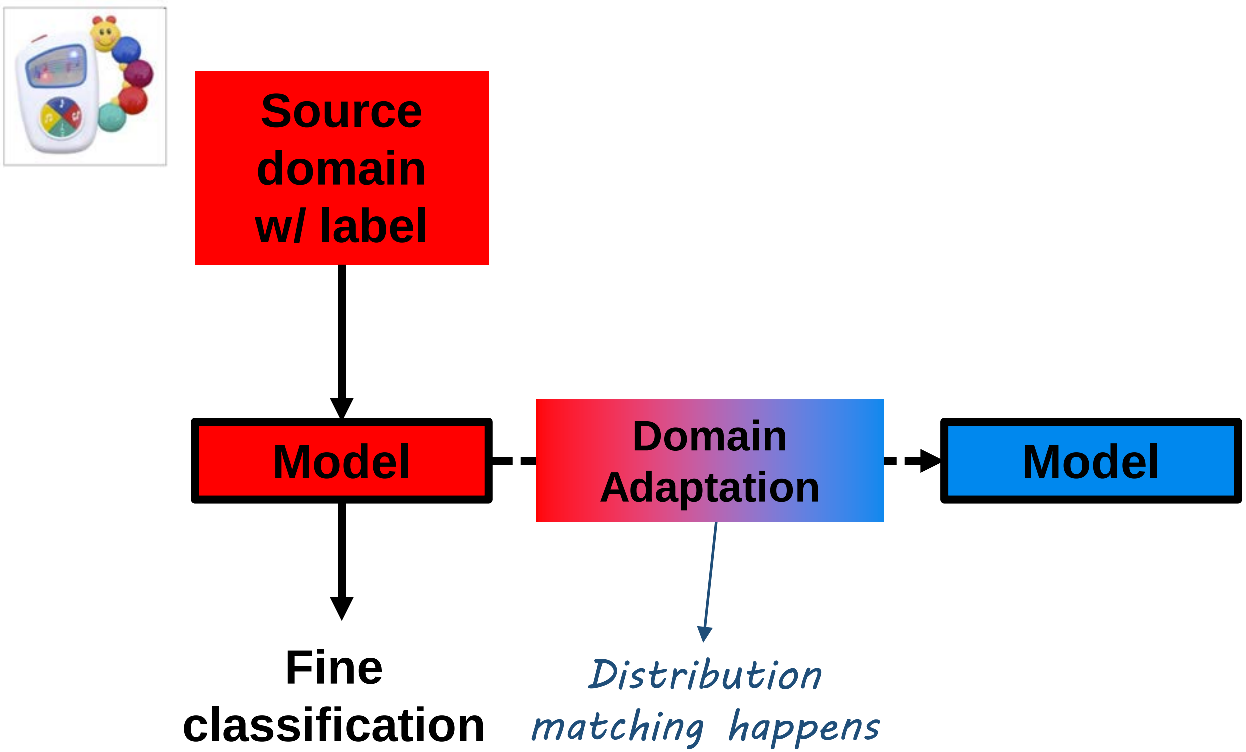
Domain Adaptation (DA)

- Adjust a model to a different <u>target domain distribution</u> starting from the <u>source domain knowledge</u>



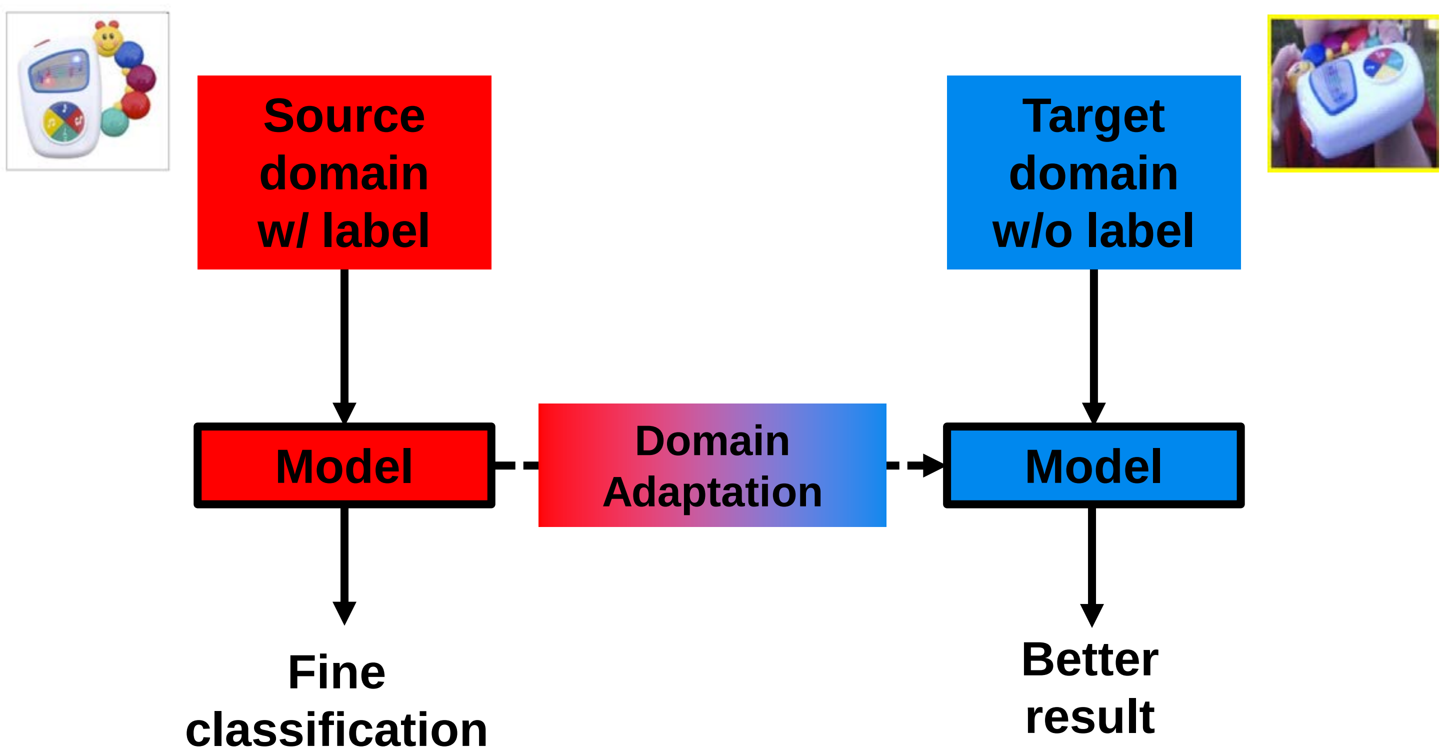
Domain Adaptation (DA)

- Adjust a model to a different <u>target domain distribution</u> starting from the <u>source domain knowledge</u>



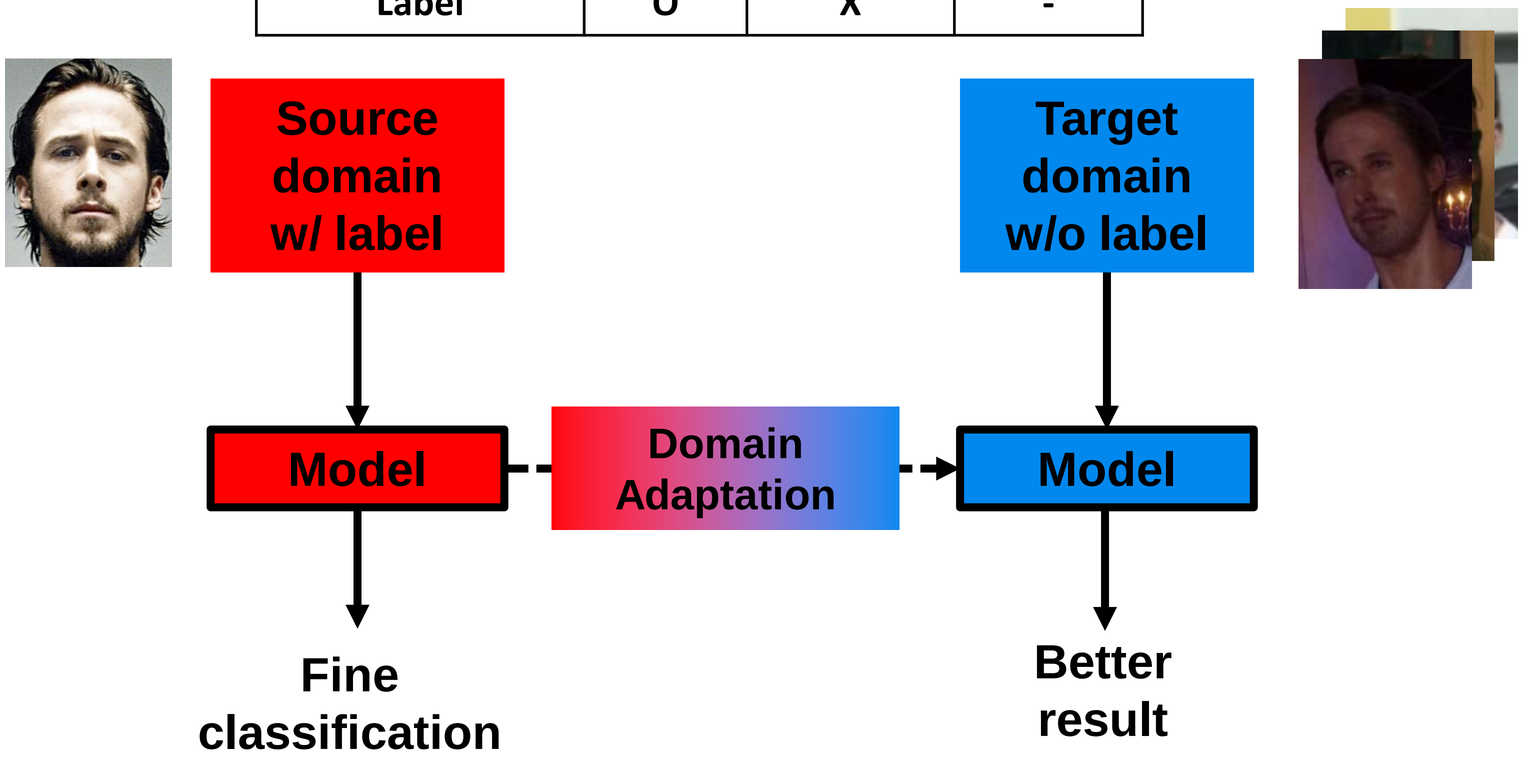
Domain Adaptation (DA)

- Adjust a model to a different target domain distribution starting from the source domain knowledge



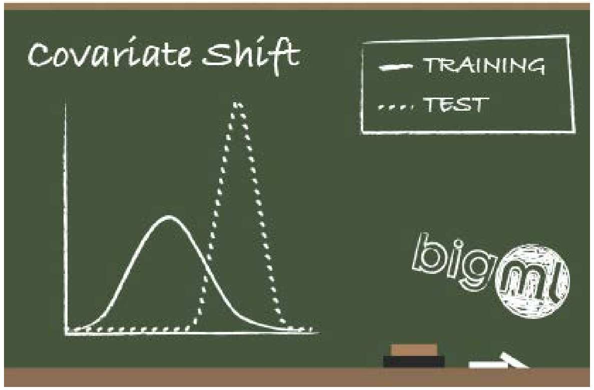
Domain Adaptation (DA)

Purpose	Train		Test
Domain	Source	Target	Target
Image condition	Stable	Unstable	Unstable
Label	O	X	-



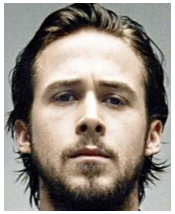
Domain Adaptation (DA)

- Basic assumptions of DA
  - samples are **abundant** in each domain
  - sample distribution of each domain is **related but different**



Lack of samples

Registration
(**Gallery**)



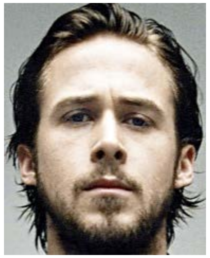
Domain Adaptation

Identification
(**Probe**)

Face synthesis

- Generate virtual samples for lack of samples.

Registration
(**Gallery**)



Synthesized
(from Gallery)

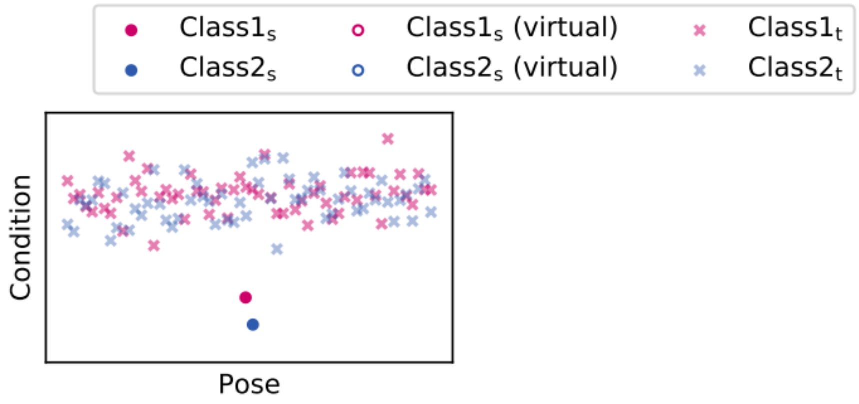
Identification
(**Probe**)

KAIST

SSPP-DAN

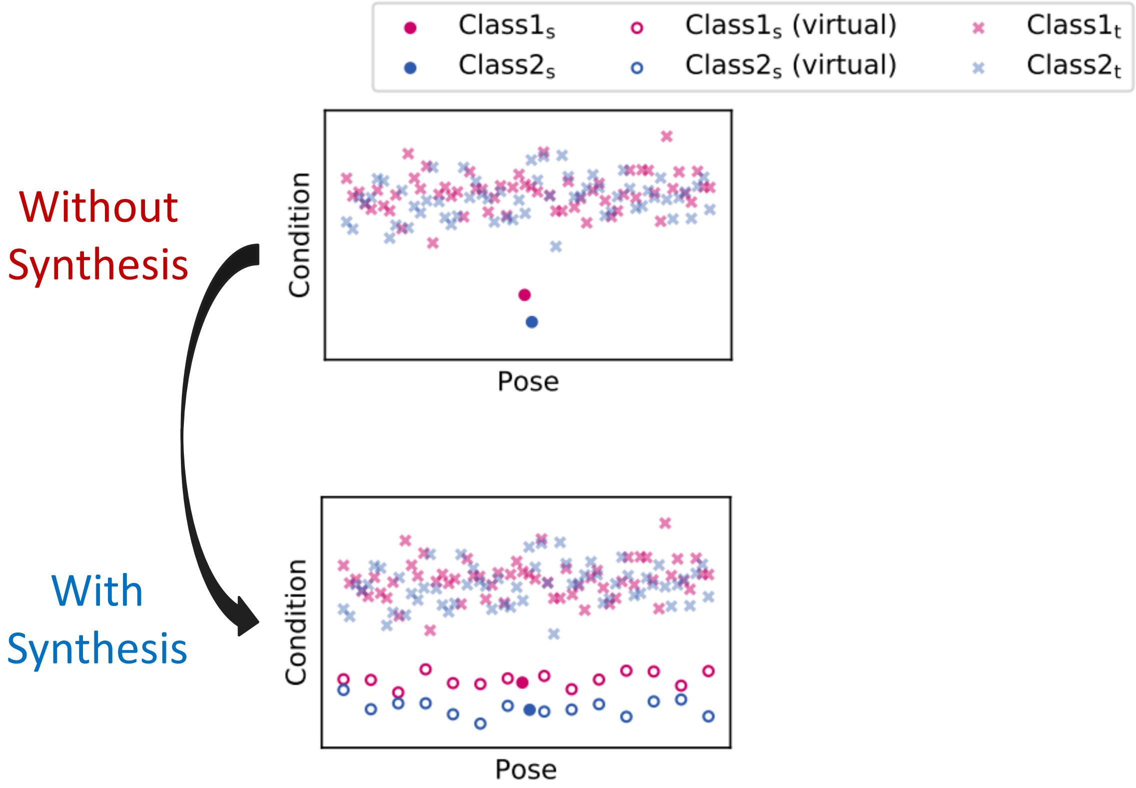
- Image synthesis

Without Synthesis



SSPP-DAN

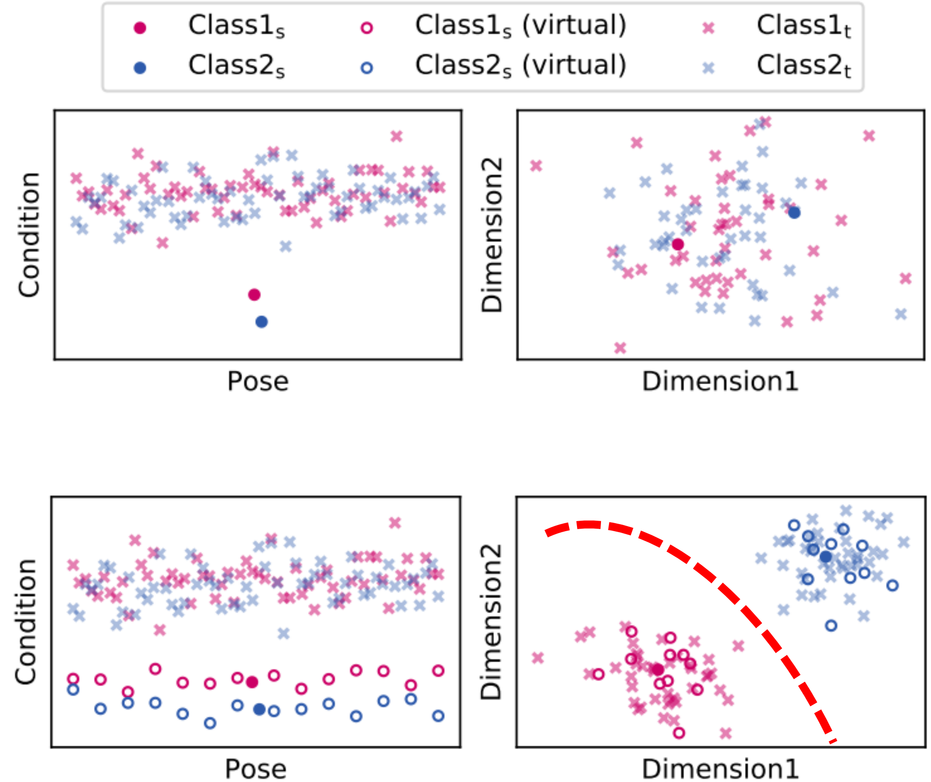
- Image synthesis
  >> **distribution of samples**



SSPP-DAN

- Image synthesis
  >> **distribution of samples** >> **Success of DA**

Without Synthesis

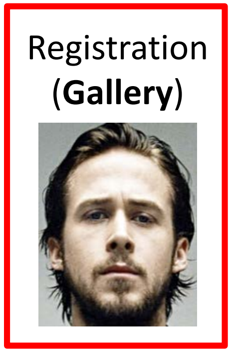


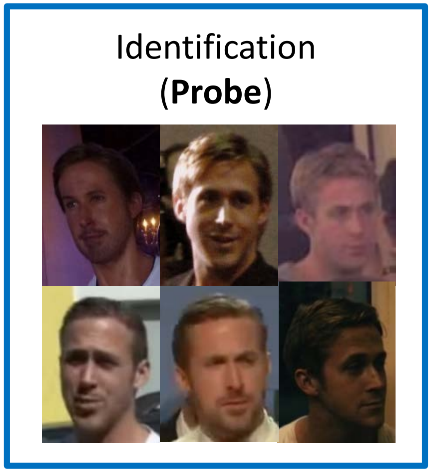
DA fails

With Synthesis

DA succeeds

Real-world SSPP Face recognition

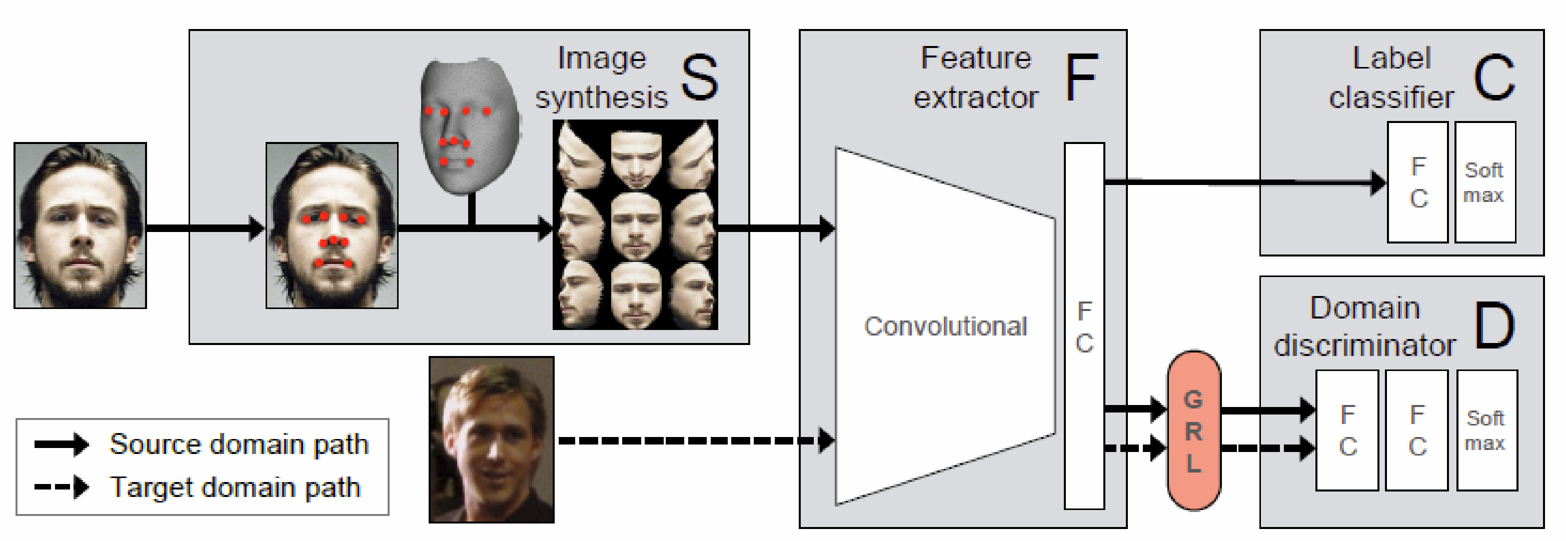




Challenges

1. **Heterogeneity of the shooting environments**
   - Gallery: stable environment
   - Probe: highly unstable environment
2. **Shortage of training samples**
   - Only one training sample per person is available

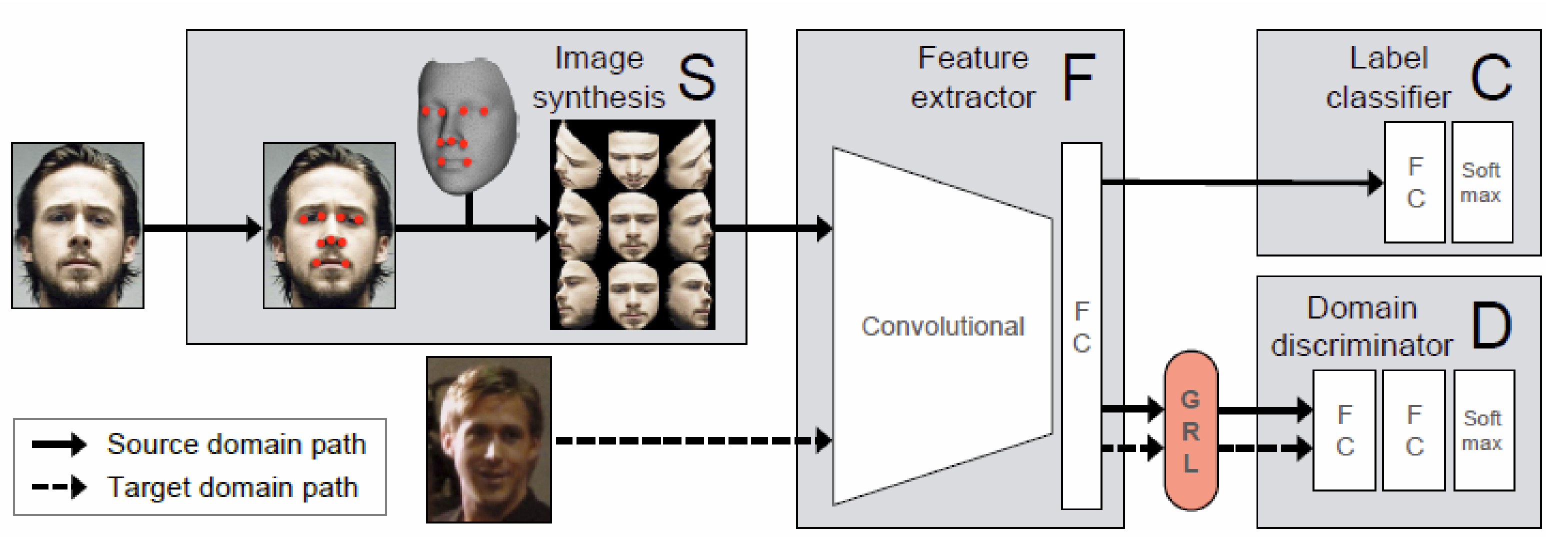
SSPP-DAN: Domain Adaptation Network for Single Sample Per Person



1. **Domain adaptation network**
   - From stable face domain (source) to unstable face domain (target)
2. **Face synthesis**
   - Generate virtual samples

KAIST

SSPP-DAN



1. **Domain adaptation network**
   with domain-adversarial training

- Feature learning
- Domain adaptation
- Classifier learning

Jointly

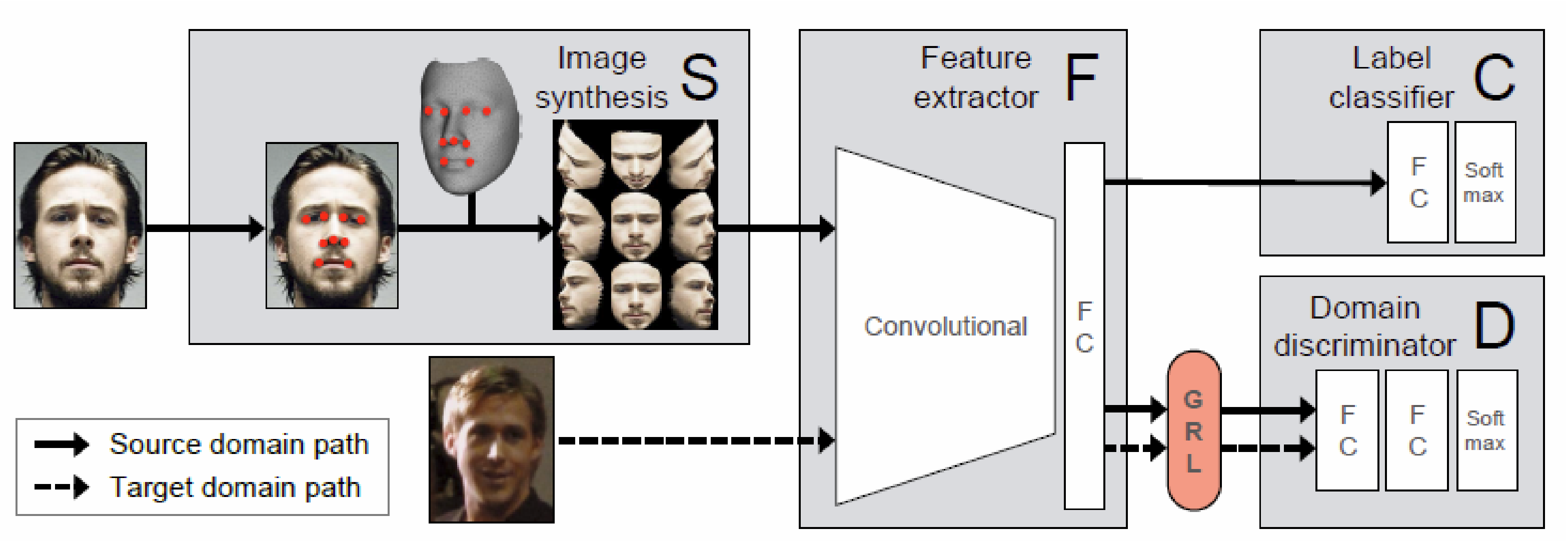
$$L_C = \sum_{i \in S} L_C^i \qquad \text{when update } \theta_C$$

$$L_D = \sum_{i \in S \cup T} L_D^i \qquad \text{when update } \theta_D$$

$$L_F = \sum_{i \in S} L_C^i - \lambda \sum_{i \in S \cup T} L_D^i \qquad \text{when update } \theta_F$$

Ganin, Yaroslav, and Victor Lempitsky. "Unsupervised domain adaptation by backpropagation." International Conference on Machine Learning. 2015.

SSPP-DAN

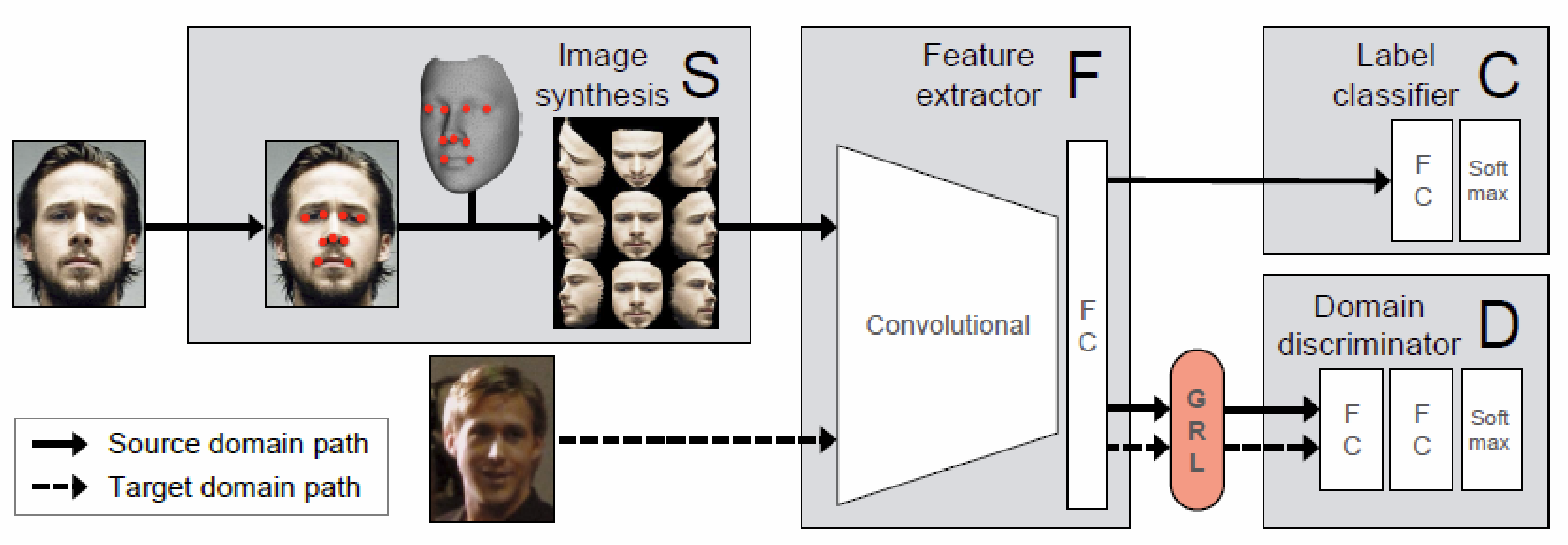


2. **Face synthesis**
   - Generate virtual samples

1) Landmark detection
   - Supervised descent method
2) 2D → 3D mapping
3) Pose estimation
4) Image synthesis
   - (yaw: -80°~+80°, pitch: -10°~40°)

Domain Adaptation

- Data Feed
  - Source (with label): frontal images + synthesized images
  - Target (without label): surveillance camera images

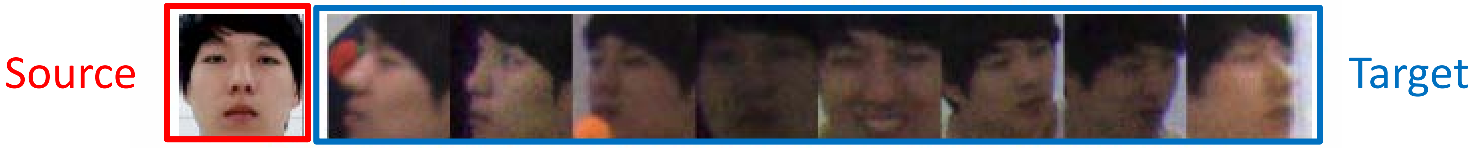


KAIST

Experiments

Heterogeneous dataset

(a) Shooting condition for the source (left) and target (center and right)



(b) Face regions from the source (leftmost) and target (the others)

Table 1: Dataset specification

Domain	Source	Target
Set	webcam	surveillance
Subjects	30	30
Samples	30	15,900
Pose	frontal	various
Condition	stable	unstable (blur, noise, illumination)

KAIST

Heterogeneous dataset

Table 2: Recognition rates (%) for different models and different training sets of the EK-LFH

	Model	Training set	Accuracy
Lower bound	Source only	S	39.22
		$S + S_v$	37.15
	DAN	$S + T$	31.11
	SSPP-DAN	$S + S_v + T$	**58.53**
Upper bound	Train on target	T_l	88.31

S: Labeled webcam T: Unlabeled surveillance
S_v: Virtual set from S T_l: Labeled surveillance

only using face synthesis

only using domain adaptation

Heterogeneous dataset

Table 2: Recognition rates (%) for different models and different training sets of the EK-LFH

Model	Training set	Accuracy
Source only	S	39.22
	S + S_v	37.15
DAN	S + T	31.11
SSPP-DAN	S + S_v + T	**58.53**
Semi DAN	S + T + T_l	67.28
Semi SSPP-DAN	S + S_v + T + T_l	**72.08**
Train on target	T_l	88.31

S: Labeled webcam T: Unlabeled surveillance
S_v: Virtual set from S T_l: Labeled surveillance

Semi: 3 samples per person from target domain are revealed

Labeled Faces in the Wild (LFW)

- Dataset summary
  - Rearranged LFW-A for SSPP face recognition
  - Gallery: 50 images for 50 people
  - Generic set: 108 subjects

Labeled Faces in the Wild (LFW)

Source

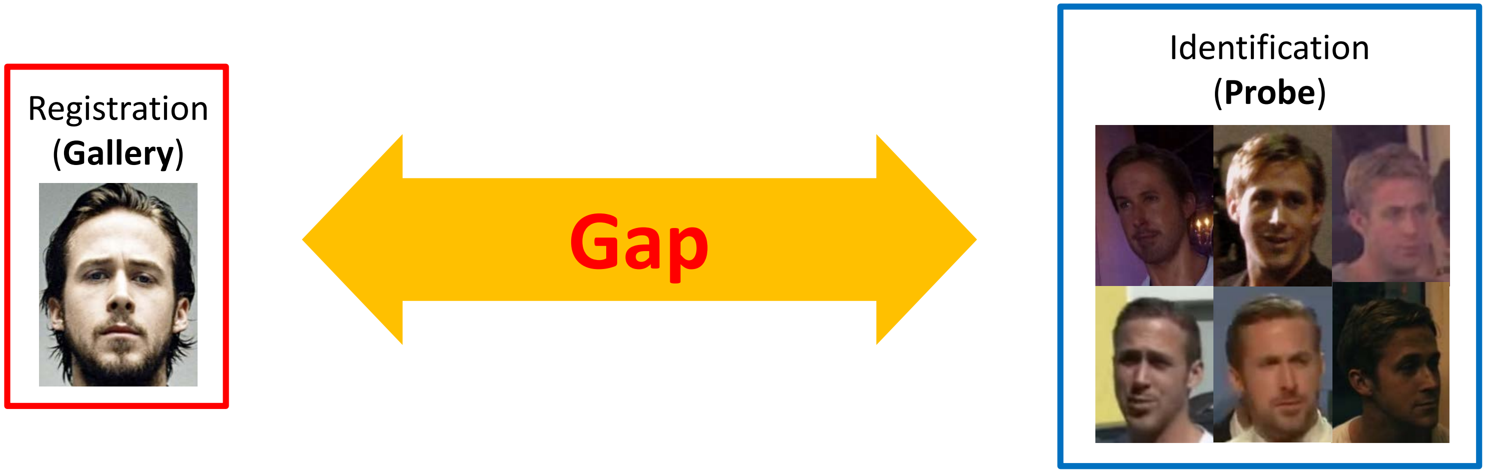
Target

- LFW for SSPP protocol
  - Gallery: 50 images for 50 people
  - Generic set: 108 subjects

Method	Accuracy	Method	Accuracy
DMMA [1]	17.8	RPR [20]	33.1
AGL [6]	19.2	DeepID [21]	70.7
SRC [4]	20.4	JCR-ACF [19]	86.0
ESRC [7]	27.3	VGG-Face [8]	96.43
LGR [22]	30.4	**Ours**	**97.91**

Deep learning

Summary

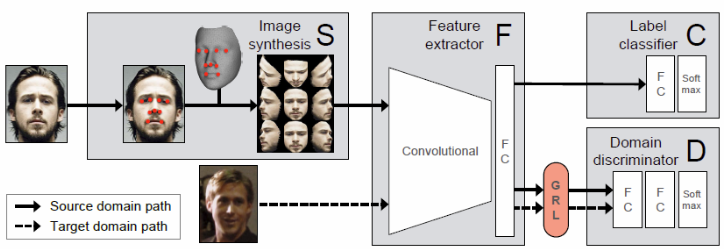


1. **Heterogeneity of the shooting environments**

2. **Shortage of training samples**

Summary

Registration
(**Gallery**)



Identification
(**Probe**)

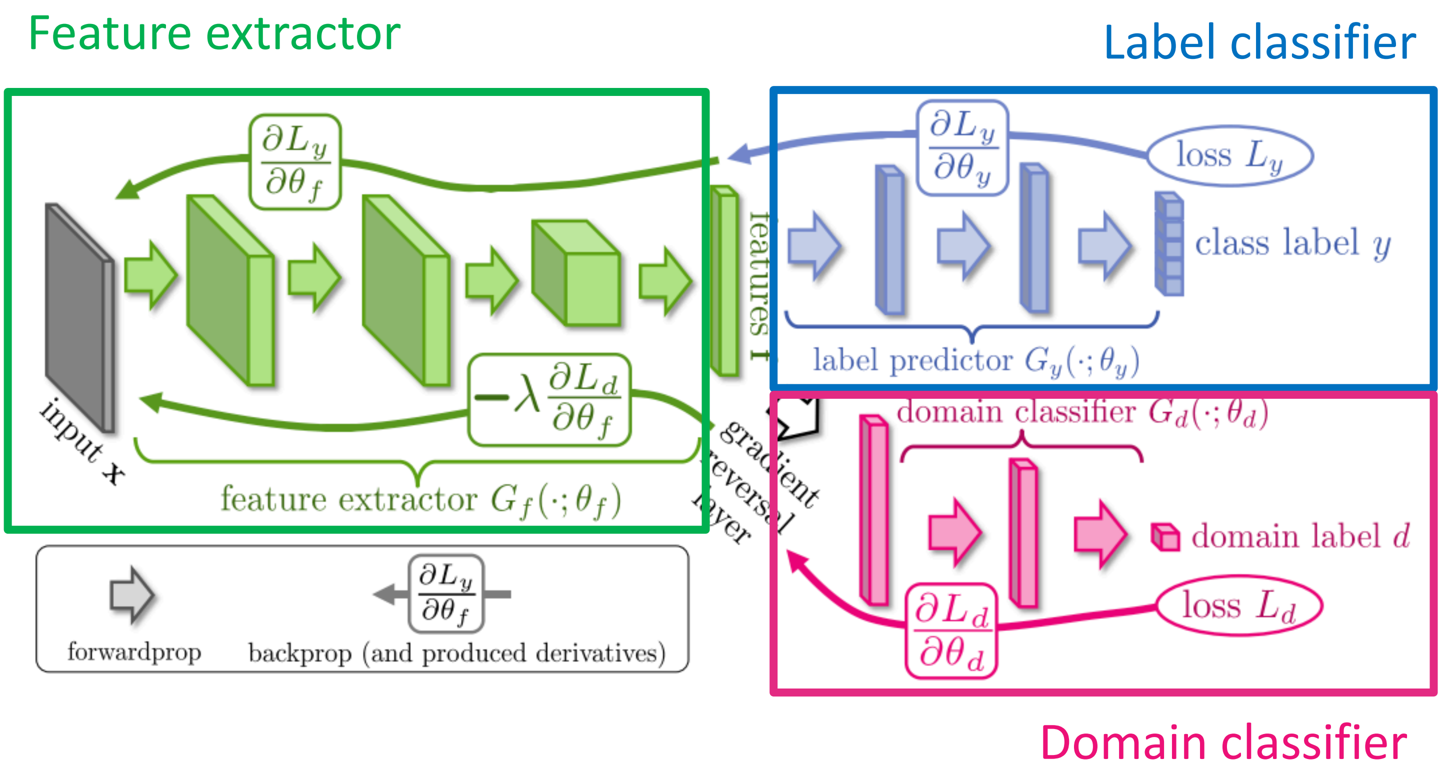
1. **Heterogeneity of the shooting environments**
   - ➢ **Domain adaptation network**
2. **Shortage of training samples**
   - ➢ **Face synthesis**

Questions?

Appendix

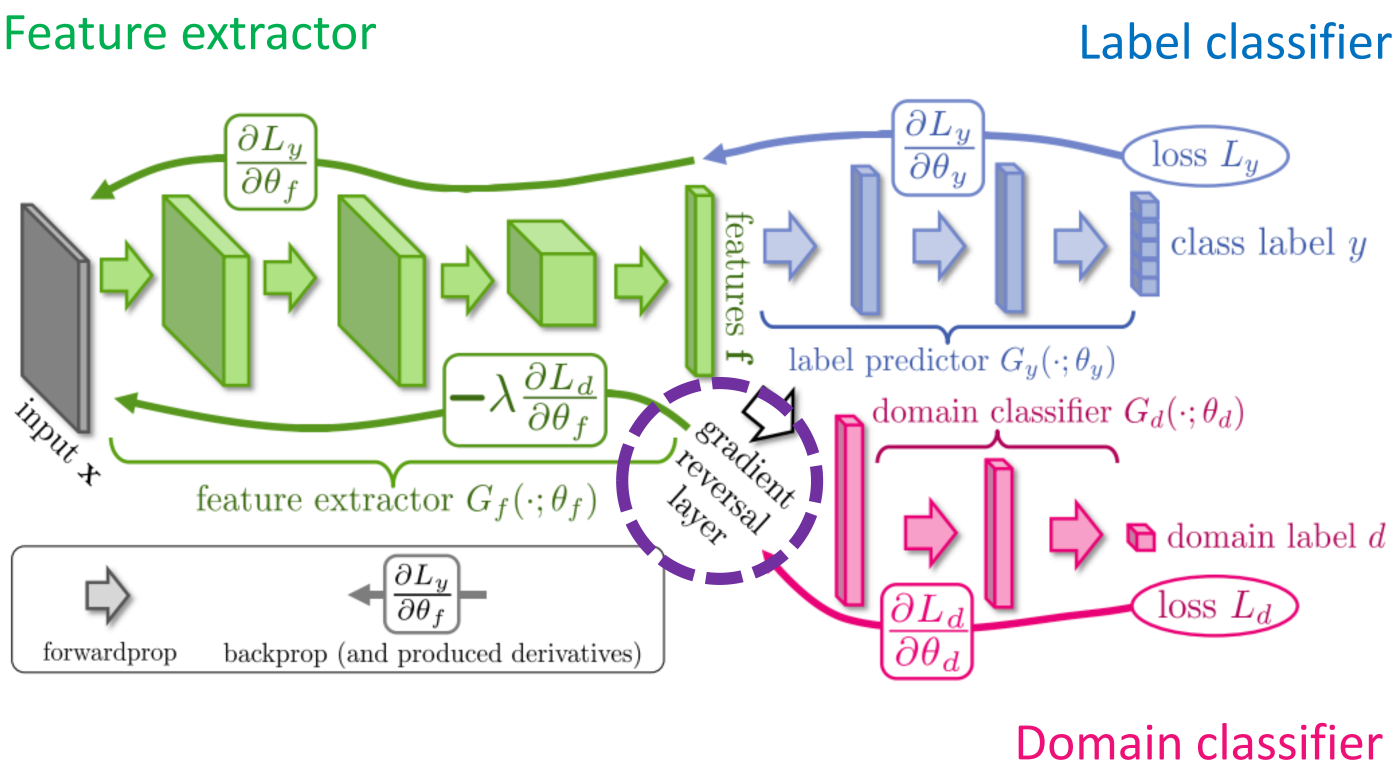
Domain Adaptation (DA)

- Domain Adversarial Network
  (Ganin, Yaroslav, and Victor Lempitsky. "Unsupervised domain adaptation by backpropagation." *ICML 2015.)*
  - Unified framework using adversarial training



Domain Adaptation (DA)

- Domain Adversarial Network
  (Ganin, Yaroslav, and Victor Lempitsky. "Unsupervised domain adaptation by backpropagation." *ICML 2015.)*
  - Unified framework using adversarial training



Domain Adaptation

- Adversarial Training
  - F is trained to fool D so that D cannot determine domain of data.
  - Gradient Reversal Layer (GRL)
    - forward: identity operation $R_\lambda(\mathbf{x}) = \mathbf{x}$
    - backward: multiply by $-\lambda$ $\dfrac{dR_\lambda}{d\mathbf{x}} = -\lambda\mathbf{I}$

- Loss for training

$$L_C = \sum_{i \in S} L_C^i \qquad \text{when update } \theta_C$$

$$L_D = \sum_{i \in S \cup T} L_D^i \qquad \text{when update } \theta_D$$

$$L_F = \sum_{i \in S} L_C^i - \lambda \sum_{i \in S \cup T} L_D^i \qquad \text{when update } \theta_F$$

L_C^i and L_D^i: loss of C and D, and
$\theta_D, \theta_F, \theta_C$: parameters of D, F, C

KAIST