How to generalize unseen domain and attack type in Face Anti Spoofing

DMQA Seminar

2024.01.05

Jung In Kim

고려대학교
KOREA UNIVERSITY

발표자 소개

❖ **김정인**

- **D**ata **M**ining & **Q**uality **A**nalytics Lab
- 석·박 통합과정 (2021.09~)

❖ **관심 연구 분야**

- Reinforcement Learning
- Self & Semi – supervised Learning

❖ **Contact**

- jungin_kim23@korea.ac.kr

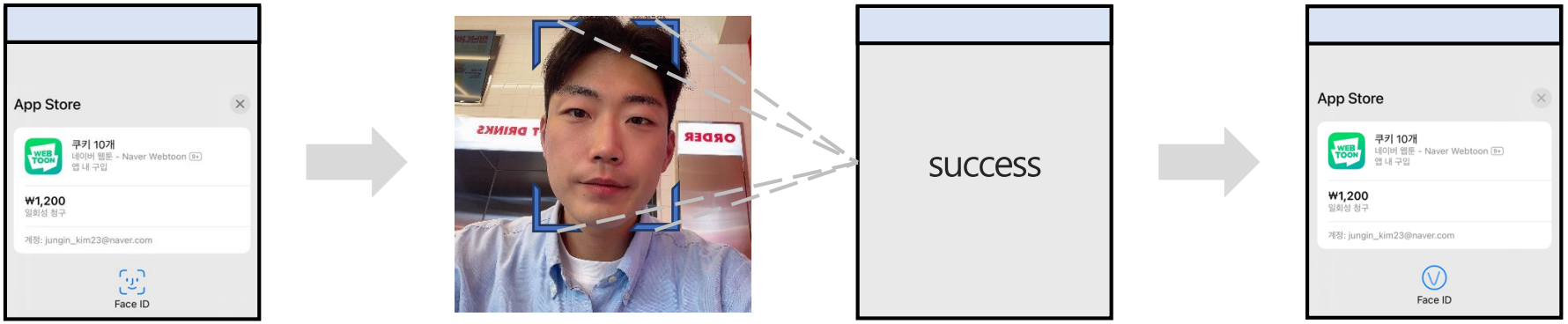
목차

Introduction to Face Anti Spoofing

Face Anti Spoofing의 중요성

❖ 얼굴 인식 적용 사례

- Check-in: 호텔, 공항 등에서 방문자가 도착하여 등록하는 과정
- Mobile payment: 스마트폰이나 태블릿과 같은 모바일 기기를 사용하여 상품이나 서비스에 대한 결제하는 방법



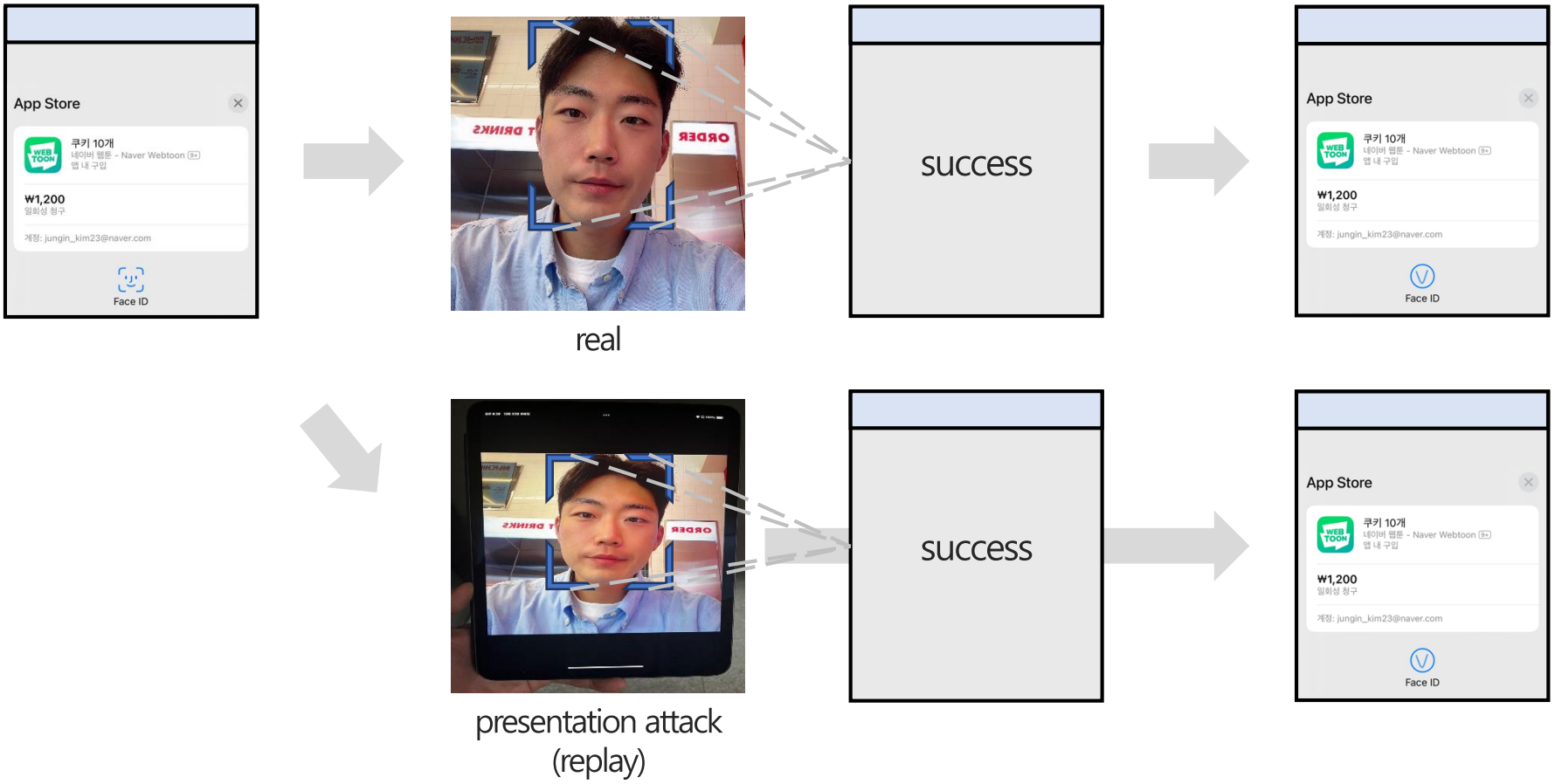
<Mobile payment>

Introduction to Face Anti Spoofing

Face Anti Spoofing의 중요성

❖ 현존하는 얼굴 인식 시스템은 presentation attack에 굉장히 취약한 문제가 존재함

- Presentation attack: 얼굴 인식 시스템을 속이기 위해 가짜 또는 조작된 데이터 (ex: print, replay, makeup, 3D-mask, etc)

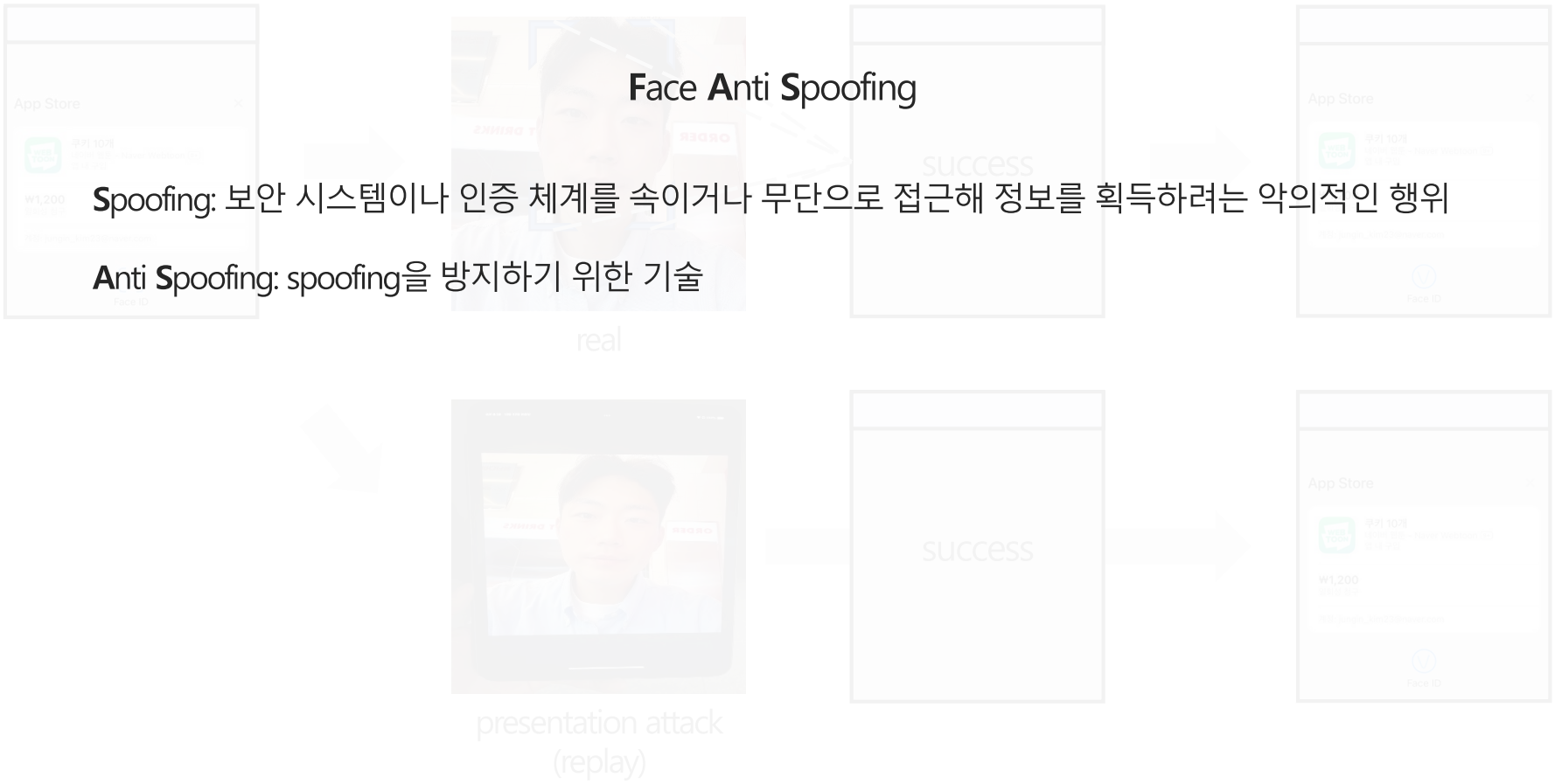


Introduction to Face Anti Spoofing

Face Anti Spoofing의 중요성

❖ 현존하는 얼굴 인식 시스템은 presentation attack에 굉장히 취약한 문제가 존재함

- Presentation attack: 얼굴 인식 시스템을 속이기 위해 가짜 또는 조작된 데이터 (ex: print, replay, makeup, 3D-mask, etc)



Introduction to Face Anti Spoofing

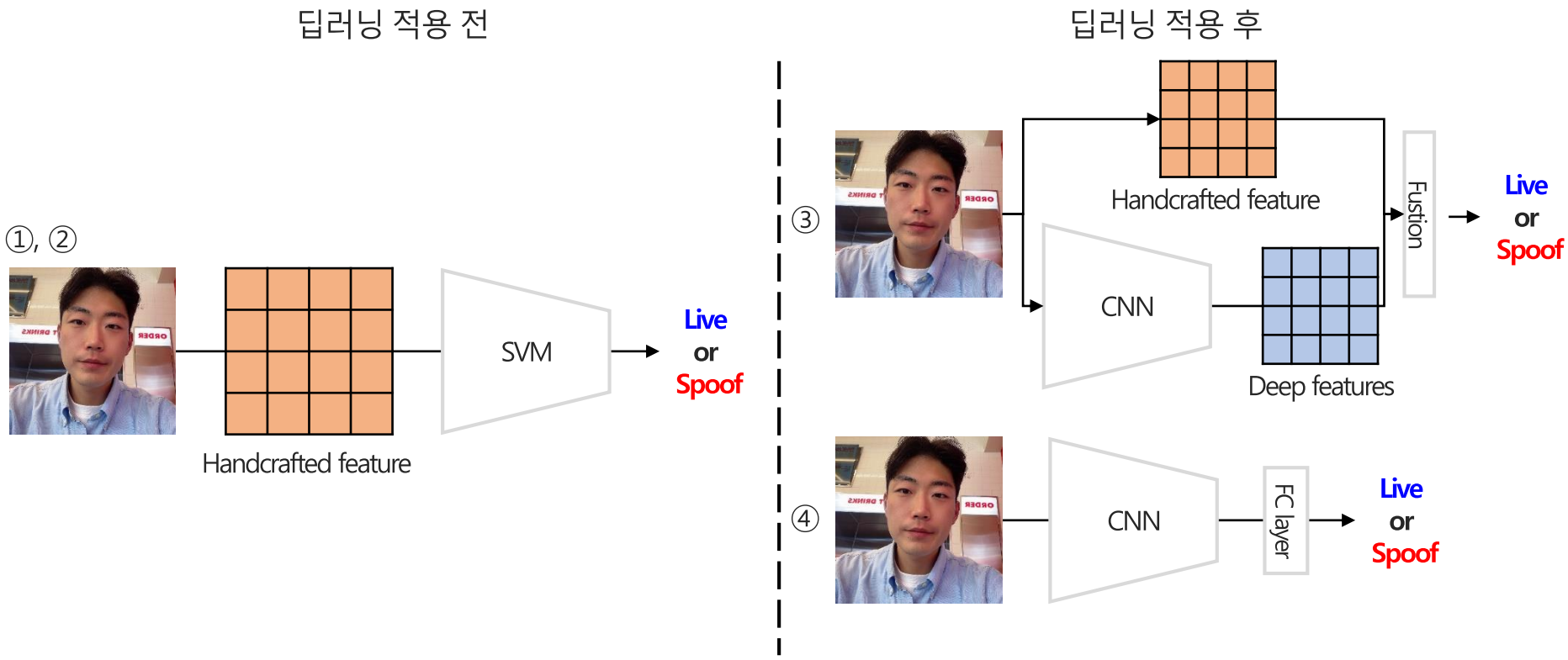
Evolution of **F**ace **A**nti **S**poofing

① 초기에는 생체 신호(눈 깜박임, 얼굴 및 머리 움직임, etc)와 수작업 특징을 기반으로 live 또는 spoof 분류

② Spoof 데이터의 특징을 추출하기 위해 고전적인 방법(LBP, SIFT, SURF, etc)을 사용하여 live 또는 spoof 분류

③ 수작업 특징과 딥러닝을 결합한 접근 방식을 사용하여 live 또는 spoof 분류

④ 딥러닝을 기반으로 live 또는 spoof 분류



Introduction to Face Anti Spoofing

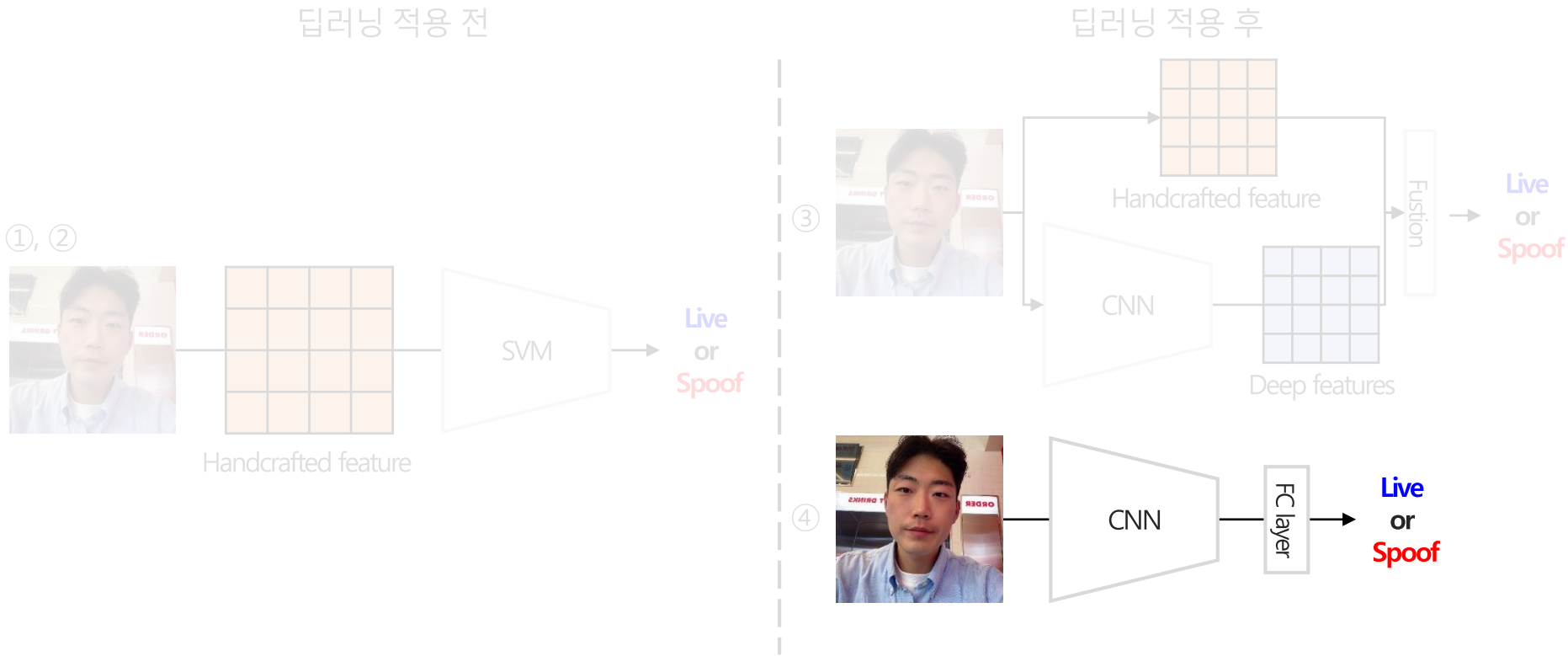
Evolution of **F**ace **A**nti **S**poofing

① 초기에는 생체 신호(눈 깜박임, 얼굴 및 머리 움직임, etc)와 수작업 특징을 기반으로 live 또는 spoof 분류

② Spoof 데이터의 특징을 추출하기 위해 고전적인 방법(LBP, SIFT, SURF, etc)을 사용하여 live 또는 spoof 분류

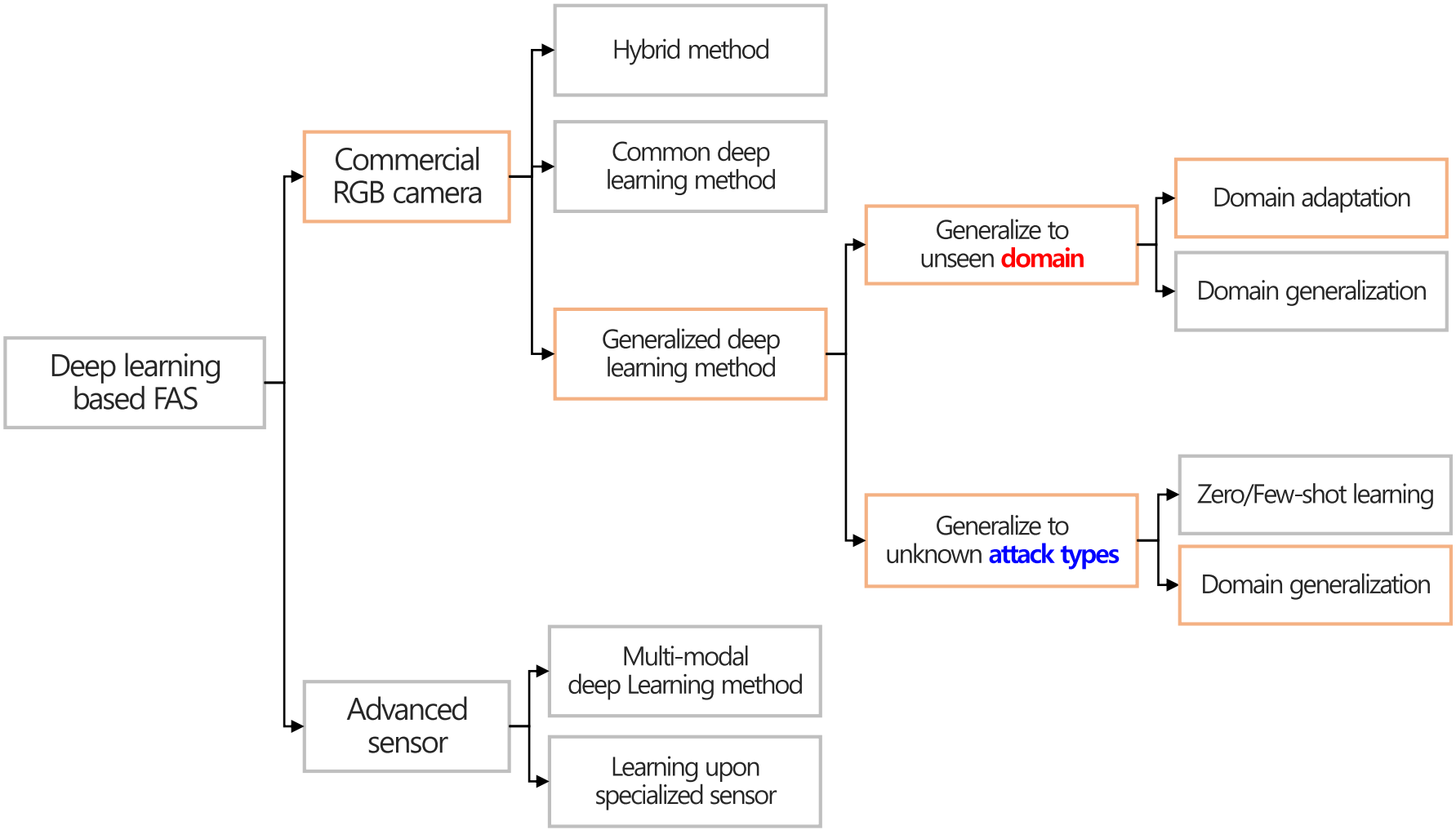
③ 수작업 특징과 딥러닝을 결합한 접근 방식을 사용하여 live 또는 spoof 분류

④ 딥러닝을 기반으로 live 또는 spoof 분류



Introduction to Face Anti Spoofing

Deep Learning based **F**ace **A**nti **S**poofing



Introduction to Face Anti Spoofing

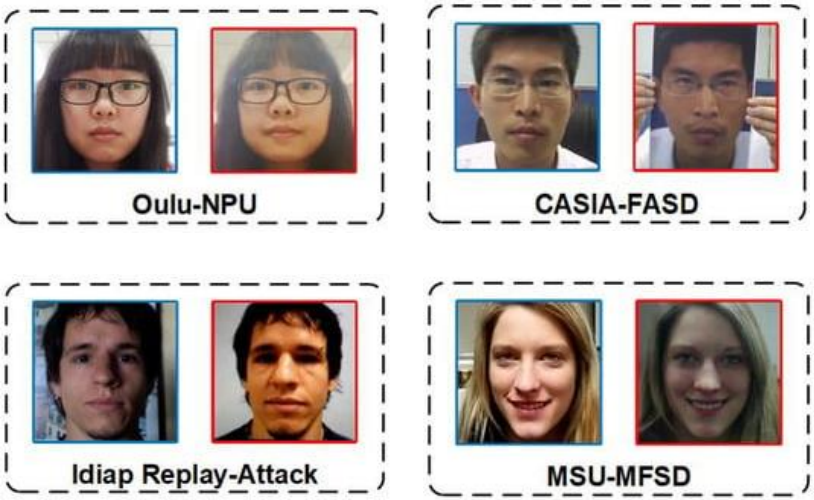
Dataset

❖ 일반화가 요구되는 부분이 크게 2가지 존재

① 보지 못했던 도메인에 대한 일반화

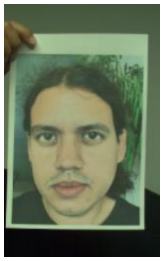
② 보지 못했던 공격 타입(attack type)에 대한 일반화

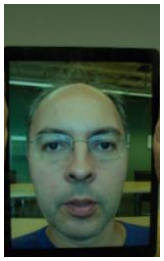
다양한 도메인



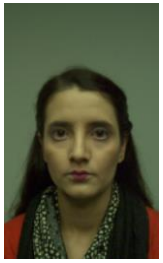
선행 연구에서 Dataset 하나를 독립적인 도메인으로 취급

공격 타입 (attack type)


Print (2D)


Replay (2D)


Paper Mask (3D)


Makeup


Resin Mask (3D)


Mannequin (3D)

Introduction to Face Anti Spoofing

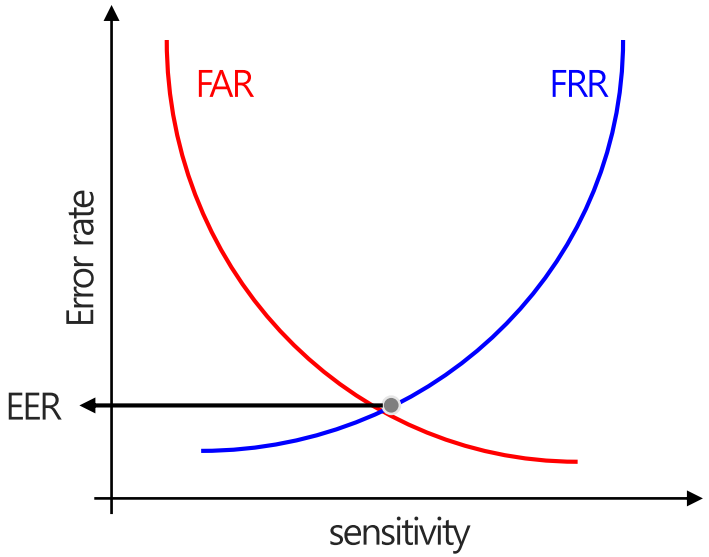
Evaluation Metrics

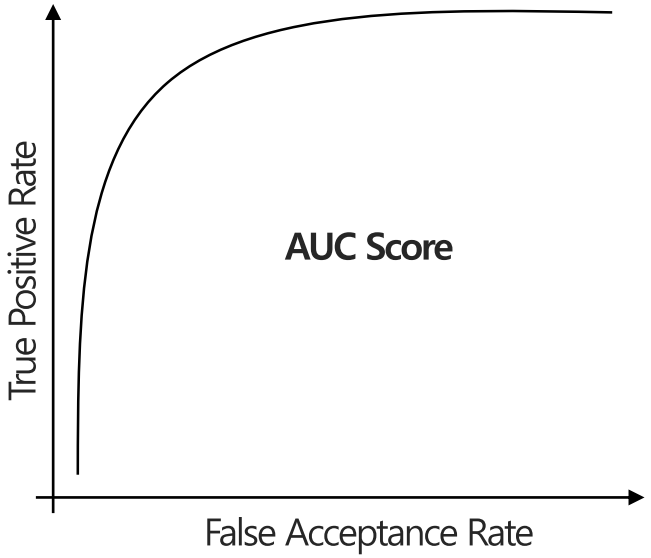
$$FAR = \frac{False\ Positive}{total\ number\ of\ spoofing\ attacks}$$

$$FRR = \frac{False\ Negative}{total\ number\ of\ living\ faces}$$

❖ 평가 지표

- 생체 인증 시스템에서 사용하고 있는 평가 지표로써, EER과 HTER은 작을 수록, AUC는 클수록 성능이 우수함
- FAR (False Acceptance Rate): 가짜 얼굴 데이터(spoof)를 실제 얼굴 데이터(live)로 잘못 인식하는 비율
- FRR (False Rejection Rate): 실제 얼굴 데이터(live)를 가짜 얼굴 데이터(spoof)로 잘못 인식하는 비율
- EER (Equal Error Rate): intra-database scenario에서 사용되는 평가 지표이며, FAR과 FRR이 같아지는 지점에서의 오류율
- HTER (Half Total Error Rate): inter-database scenario에서 사용되는 평가 지표이며, FAR과 FRR의 평균 값으로 표현됨
- AUC (Area Under Curve): ROC 곡선 아래의 면적을 측정한 값





Methods of Face Anti Spoofing

Generalize to unseen domain – domain adaptation

❖ Progressive Transfer Learning for Face Anti-Spoofing (2021, IEEE Transactions on image preprocessing)

- IEEE, senior member로 선정된 연구원들이 연구하였으며, 1월 5일 기준 5회 인용
- 해당 연구의 기여점은 크게 3가지로 요약될 수 있음
  - ① 적은 양의 레이블 데이터를 사용하는 **semi-supervised learning 기반 framework** 제안
  - ② 보지 못했던 도메인 간 차이를 줄이기 위해 **adaptive transfer mechanism** 제안
  - ③ 강건하고 신뢰할 수 있는 pseudo label을 선정하기 위해 **temporal confidence consistency mechanism**을 제안

IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 30, 2021

Progressive Transfer Learning for Face Anti-Spoofing

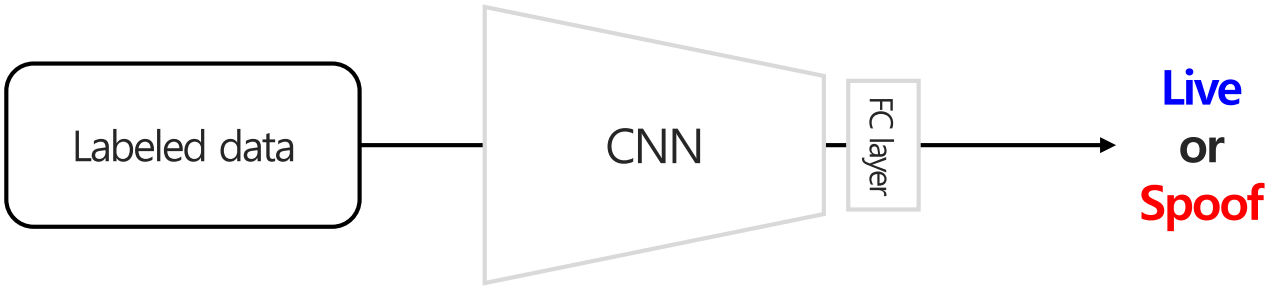
Ruijie Quan, Yu Wu, Xin Yu, and Yi Yang, *Senior Member, IEEE*

Methods of Face Anti Spoofing

Progressive Transfer Learning for Face Anti-Spoofing

❖ 연구배경 및 문제 상황

- FAS 관련 연구가 지도 학습을 기반으로 진행
  - ✓ 좋은 성능의 모델을 학습하기 위해서 많은 수의 레이블이 달린 spoof 데이터가 요구됨
  - ✓ 하지만, 실제 데이터와 spoof 데이터 간에 높은 유사성 때문에 정확한 레이블을 지정하는 것이 어려움
  - ✓ 또한, 새로운 유형의 spoof 데이터를 사전에 얻는 것이 어렵고 신속하게 새로운 유형의 spoof에 대응하기 어려움

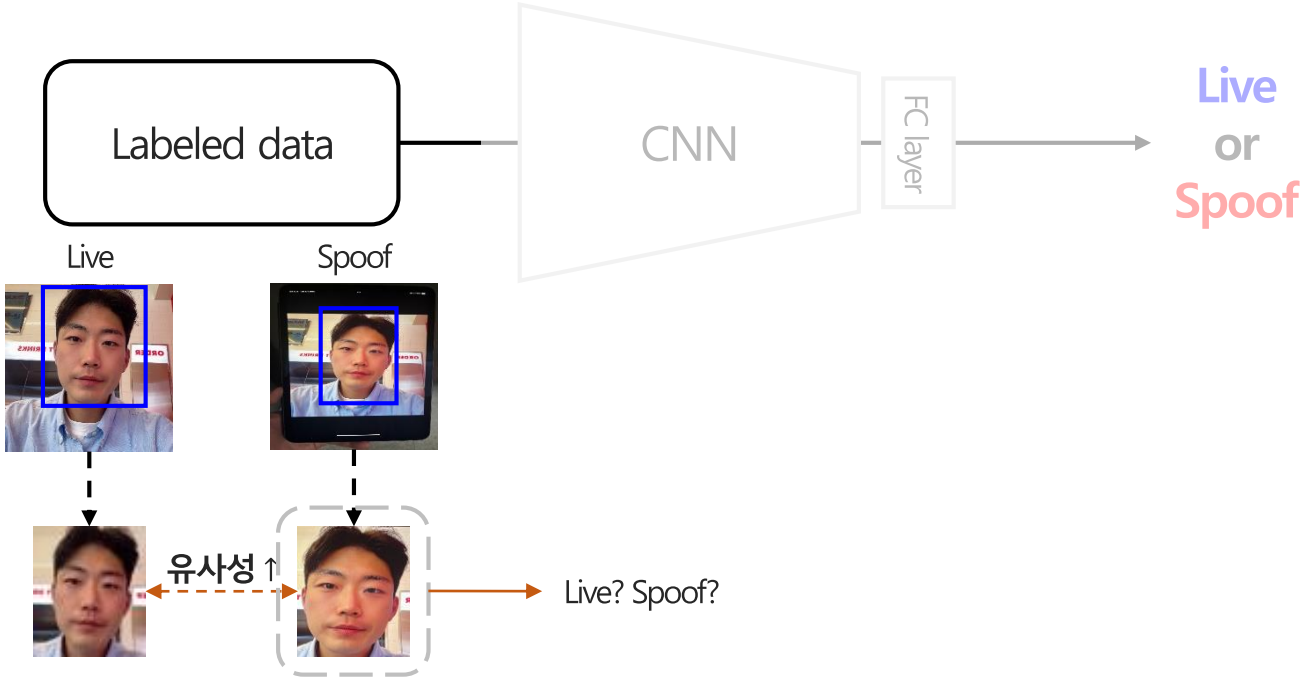


Methods of Face Anti Spoofing

Progressive Transfer Learning for Face Anti-Spoofing

❖ 연구배경 및 문제 상황

- FAS 관련 연구가 지도 학습을 기반으로 진행
  - ✓ 좋은 성능의 모델을 학습하기 위해서 **많은 수의 레이블이 달린 spoof 데이터**가 요구됨
  - ✓ 하지만, 실제 데이터와 spoof 데이터 간에 **높은 유사성** 때문에 정확한 레이블을 지정하는 것이 어려움
  - ✓ 또한, 새로운 유형의 spoof 데이터를 사전에 얻는 것이 어렵고 신속하게 새로운 유형의 spoof에 대응하기 어려움

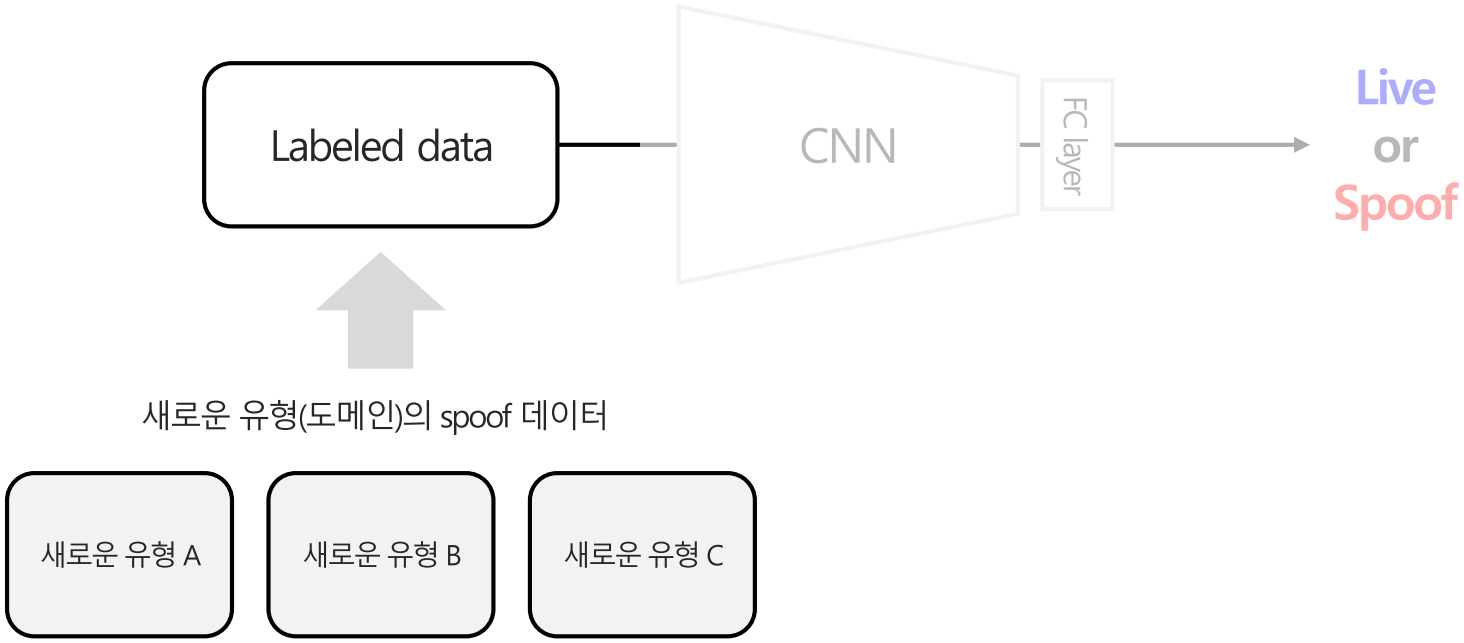


Methods of Face Anti Spoofing

Progressive Transfer Learning for Face Anti-Spoofing

❖ 연구배경 및 문제 상황

- FAS 관련 연구가 지도 학습을 기반으로 진행
  - ✓ 좋은 성능의 모델을 학습하기 위해서 많은 수의 레이블이 달린 spoof 데이터가 요구됨
  - ✓ 하지만, 실제 데이터와 spoof 데이터 간에 높은 유사성 때문에 정확한 레이블을 지정하는 것이 어려움
  - ✓ 또한, 새로운 유형의 spoof 데이터를 **사전에** 얻는 것이 어렵고 신속하게 **새로운 유형의 spoof에 대응**하기 어려움

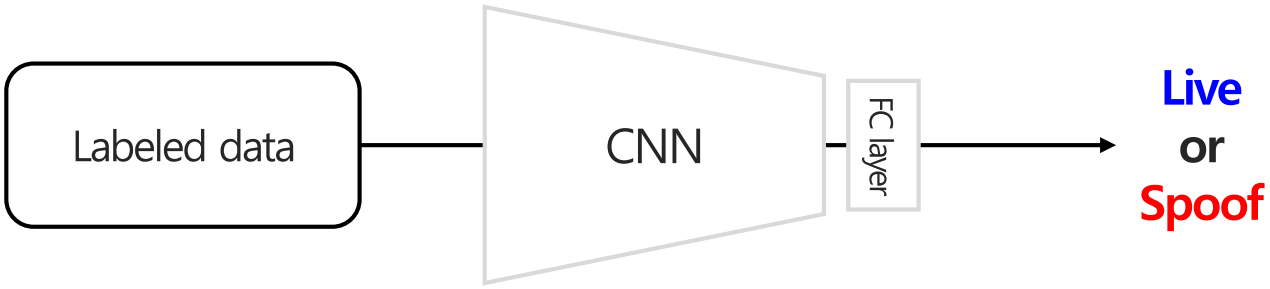


Methods of Face Anti Spoofing

Progressive Transfer Learning for Face Anti-Spoofing

❖ 연구배경 및 문제 상황

- FAS 관련 연구가 지도 학습을 기반으로 진행
  - ① 좋은 성능의 모델을 학습하기 위해서 많은 수의 레이블이 달린 spoof 데이터가 요구됨
  - ② 하지만, 실제 데이터와 spoof 데이터 간에 높은 유사성 때문에 정확한 레이블을 지정하는 것이 어려움
  - ③ 또한, 새로운 유형의 spoof 데이터를 사전에 얻는 것이 어렵고 신속하게 새로운 유형의 spoof에 대응하기 어려움
- 이러한 문제를 개선하기 위해, **semi-supervised learning 기반 framework(①,②)**와 **adaptive transfer mechanism(③)** 제안



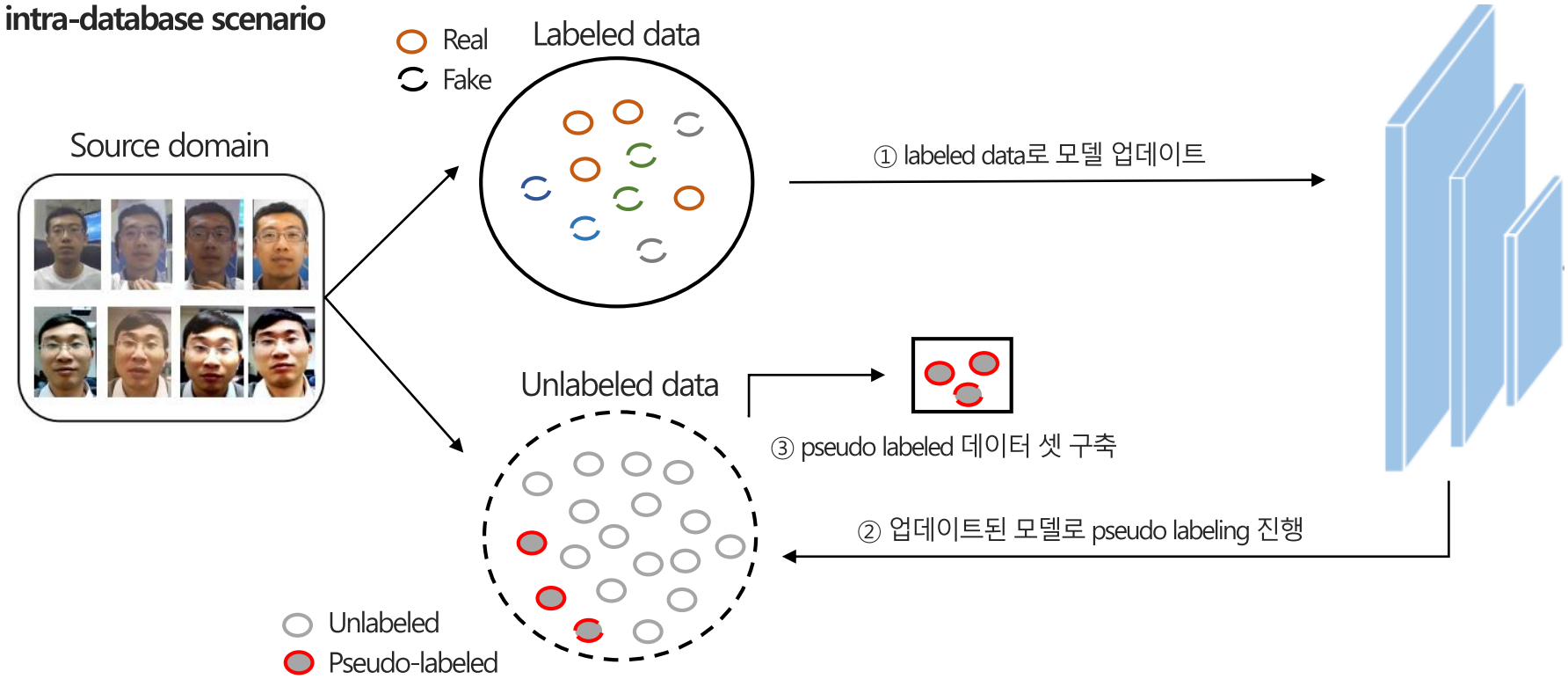
Methods of Face Anti Spoofing

Progressive Transfer Learning for Face Anti-Spoofing

❖ 제안 방법론 – progressive learning (semi-supervised learning based framework)

① Progressive learning (PL): 다수의 레이블 데이터를 얻기 어려운 문제를 해결하기 위한 실시간 학습 방법

② Adaptive Transfer (AT): 새로운 유형의 공격에 신속하게 대응하기 위한 방법



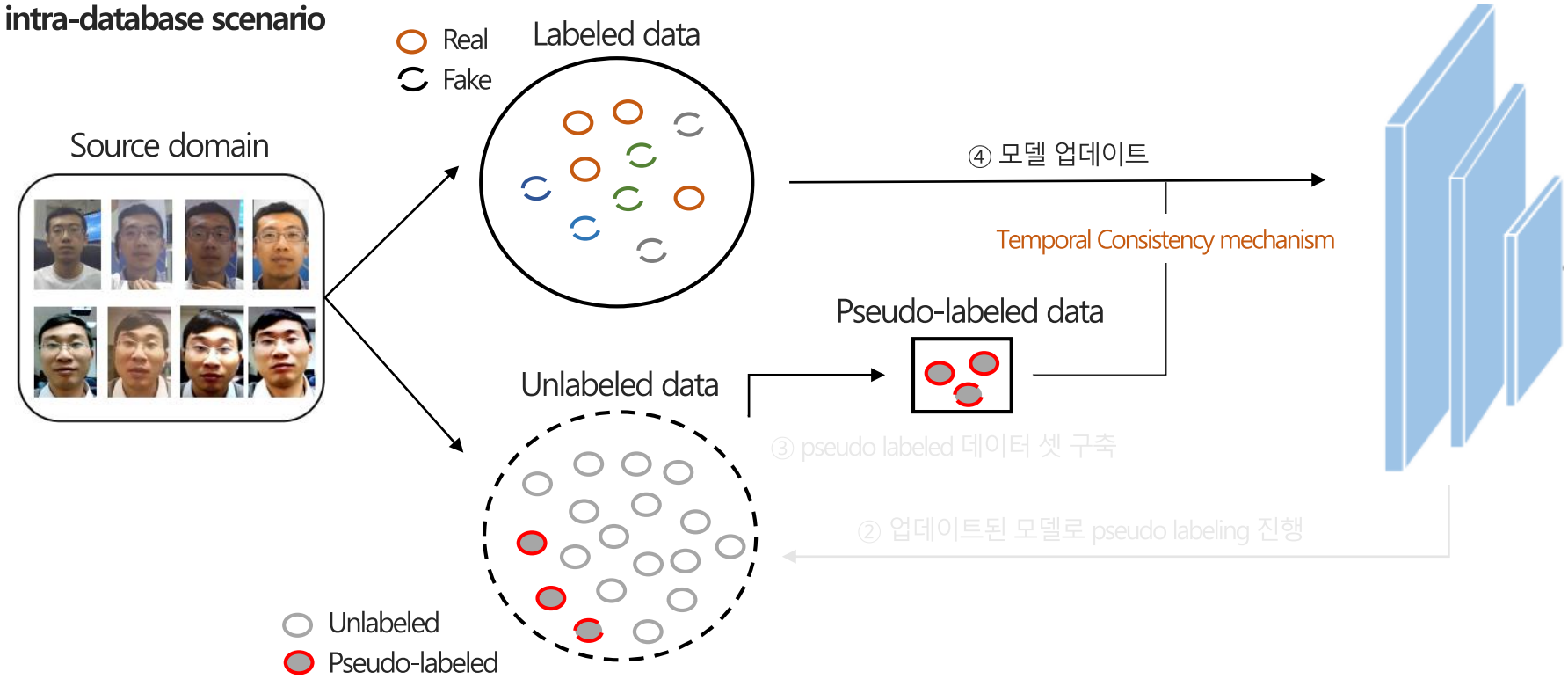
Methods of Face Anti Spoofing

Progressive Transfer Learning for Face Anti-Spoofing

❖ 제안 방법론 – progressive learning (semi-supervised learning based framework)

① Progressive learning (PL): 다수의 레이블 데이터를 얻기 어려운 문제를 해결하기 위한 실시간 학습 방법

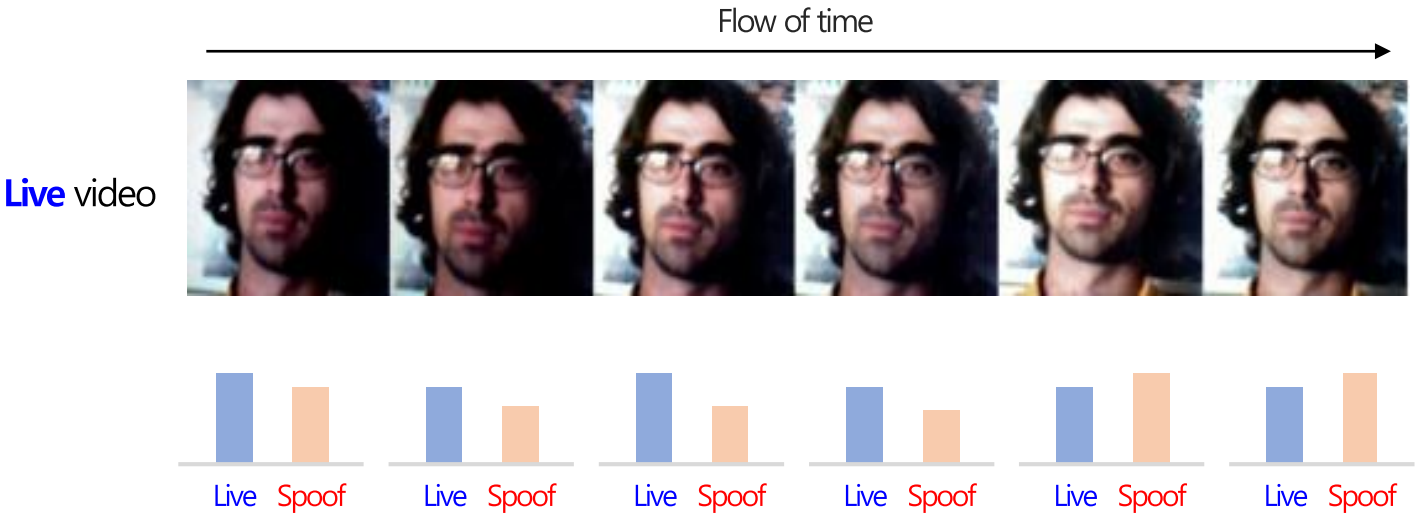
② Adaptive Transfer (AT): 새로운 유형의 공격에 신속하게 대응하기 위한 방법



Methods of Face Anti Spoofing

Progressive Transfer Learning for Face Anti-Spoofing

❖ 제안 방법론 – Temporal Consistency mechanism

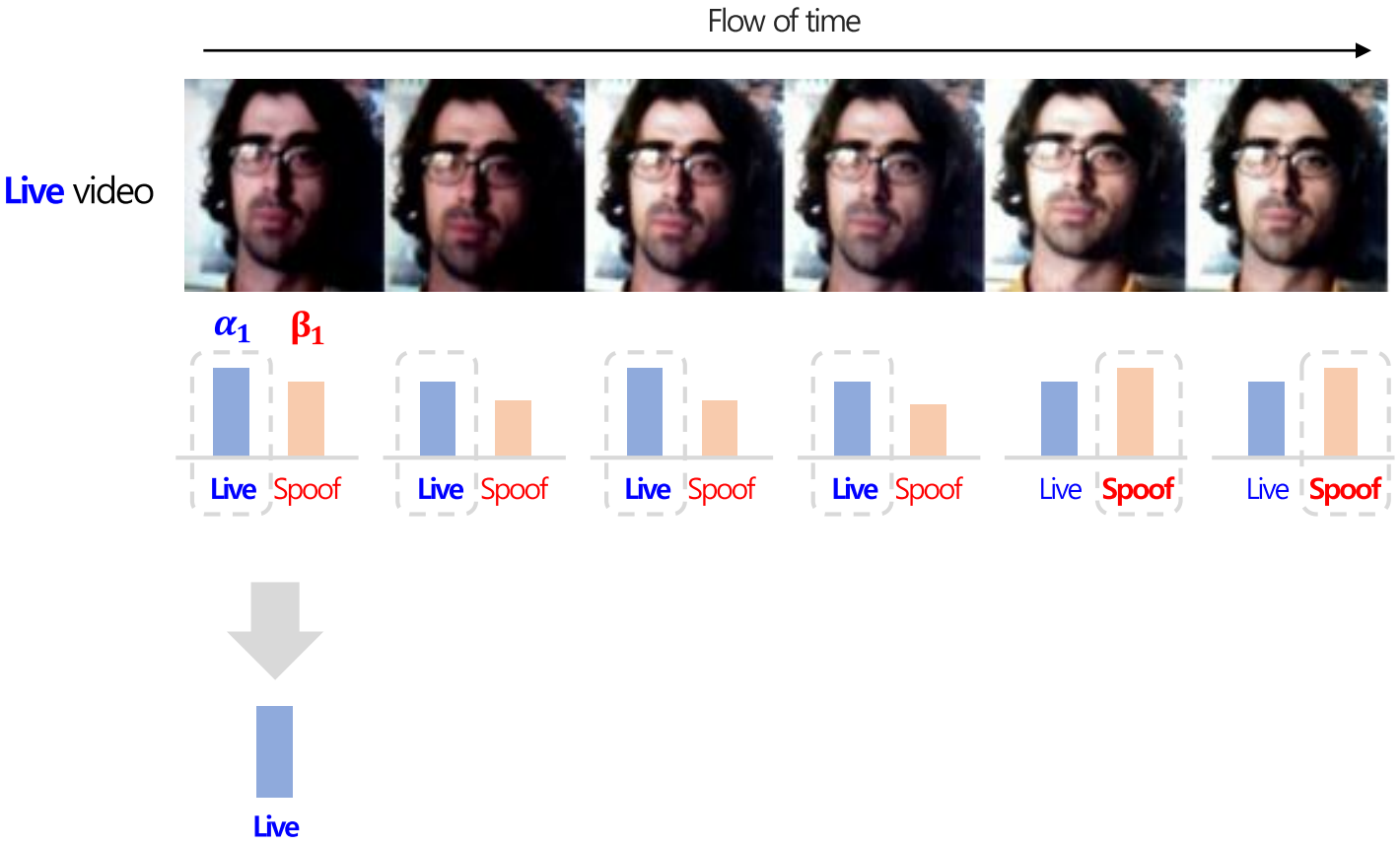


- https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9387164

Methods of Face Anti Spoofing

Progressive Transfer Learning for Face Anti-Spoofing

❖ 제안 방법론 – Temporal Consistency mechanism



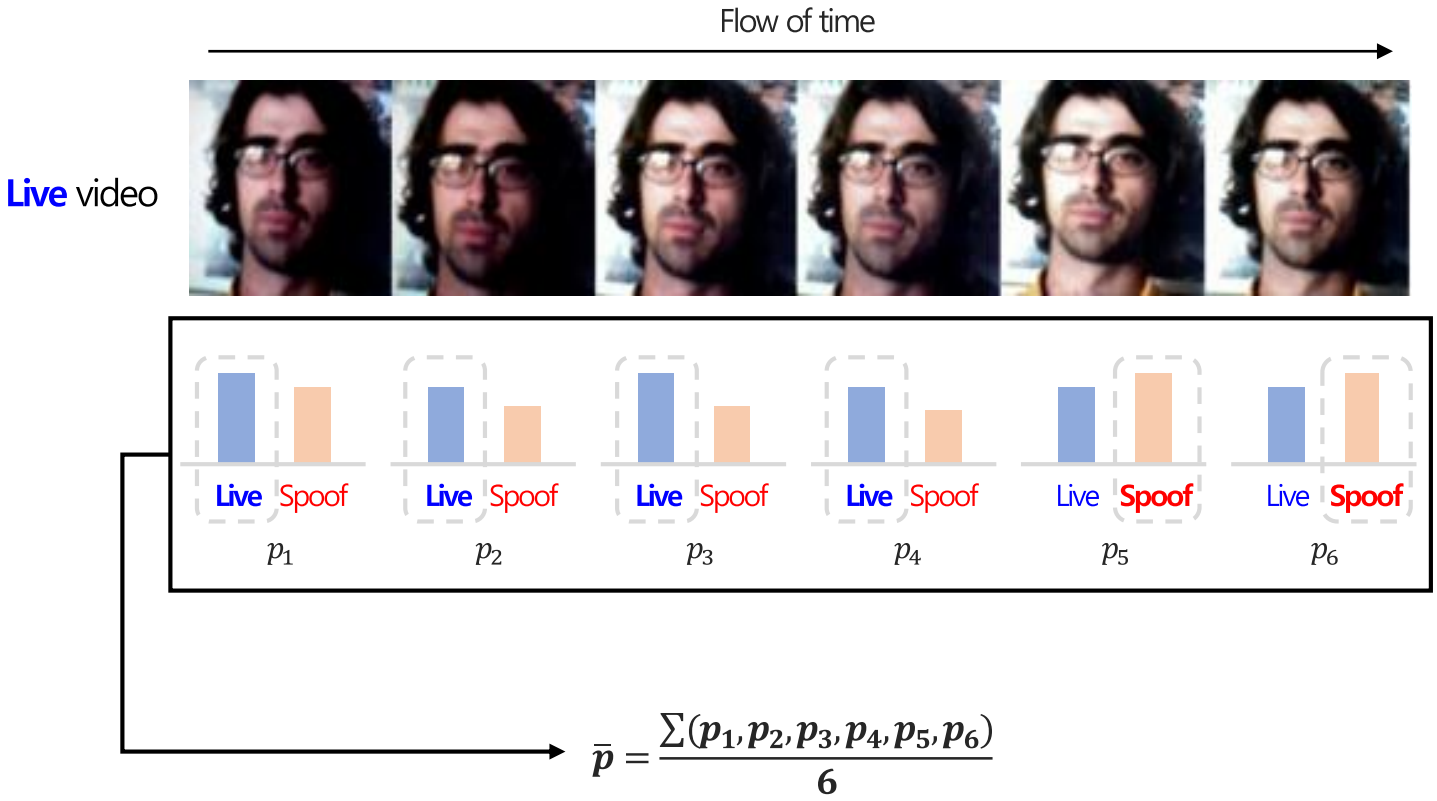
$$p_1 = \max(\alpha_i, \beta_i)$$

- https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9387164

Methods of Face Anti Spoofing

Progressive Transfer Learning for Face Anti-Spoofing

❖ 제안 방법론 – Temporal Consistency mechanism

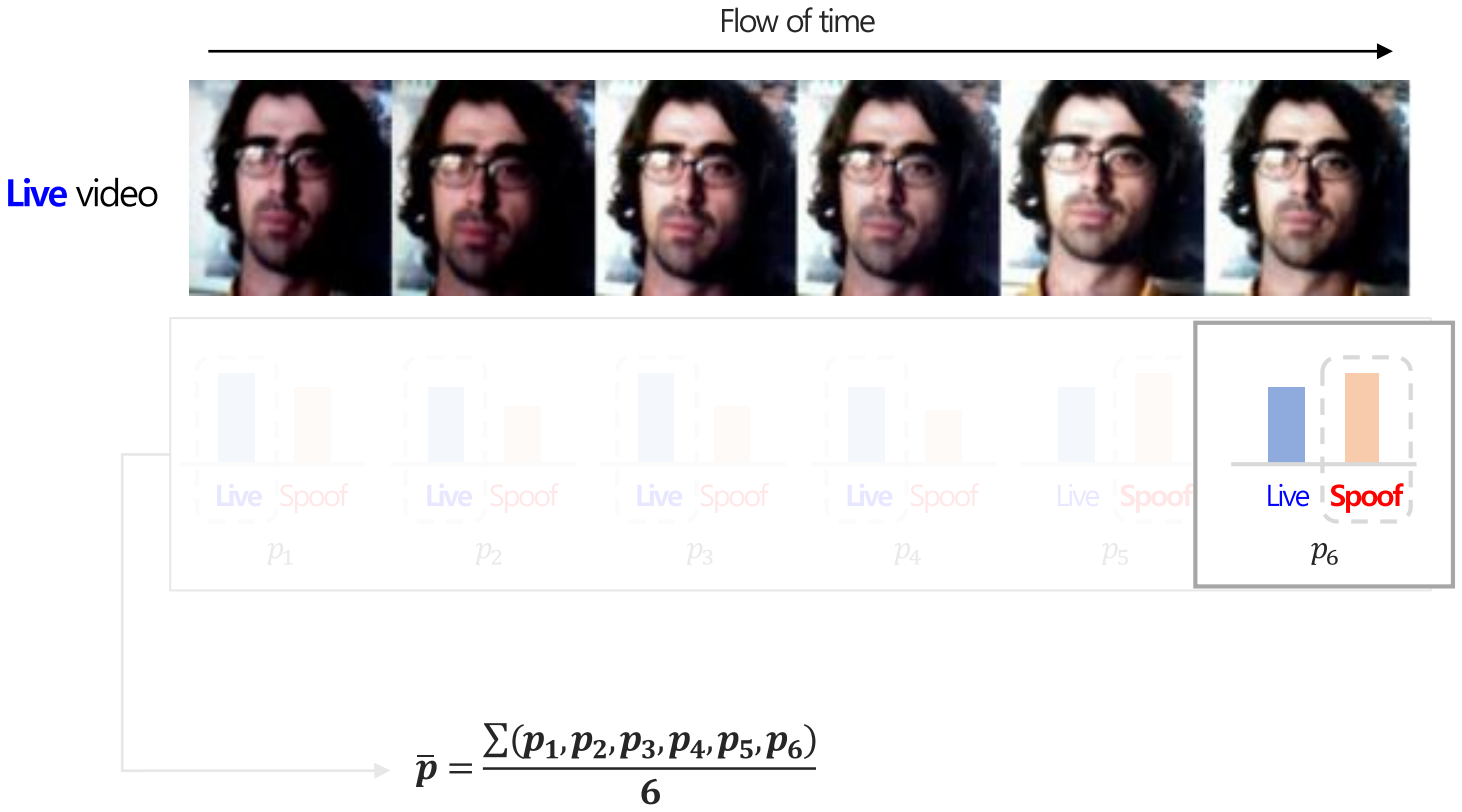


- https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9387164

Methods of Face Anti Spoofing

Progressive Transfer Learning for Face Anti-Spoofing

❖ 제안 방법론 – Temporal Consistency mechanism



- https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9387164

Methods of Face Anti Spoofing

$$C_t = \sigma_t \cdot N$$
$$(\sigma_t = k \cdot t)$$

Progressive Transfer Learning for Face Anti-Spoofing

❖ 제안 방법론 – Temporal Consistency mechanism

- 내림차순으로 정렬한 후, 사용자가 설정할 수 있는 임계 값(μ)보다 큰 데이터에만 사용함
- 매 iteration마다 top-C_t개 만큼 unlabeled data를 샘플링 진행

Flow of time

Live video

내림차순 $p_2 = 0.8$ $p_1 = 0.6$ $p_3 = 0.4$ $p_5 = 0.2$ $p_4 = 0.1$ $p_6 = 0.05$

$C_t = 0.3$ $p_2 = 0.8$ $p_1 = 0.6$ $p_3 = 0.4$ $p_5 = 0.2$ $p_4 = 0.1$ $p_6 = 0.05$

$\mu = 0.3$ $p_2 = 0.8$ $p_1 = 0.6$ $p_3 = 0.4$ $p_5 = 0.2$ $p_4 = 0.1$ $p_6 = 0.05$

- https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9387164

Methods of Face Anti Spoofing

Progressive Transfer Learning for Face Anti-Spoofing

❖ Experiment – intra database

- 지도 학습 접근 방식 모델과 비교하였을 때, 50개 레이블 데이터만을 사용하여도 더 우수한 성능을 보임

Method	REPLAY-ATTACK	CASIA-FASD	MSU-MFSD
	EER(%) ↓	EER (%) ↓	EER (%) ↓
With full labeled samples			
LBP+LDA [11]	18.25	21.01	-
IQA [37]	-	32.46	-
CDD [13]	9.75	11.85	-
IDA [32]	-	12.97	8.58
Patch-CNN [38]	0.72	4.44	-
Color [39]	0.42	2.17	4.9
GFA-CNN [40]	0.30	8.3	7.5
S-CNN	0.28	0.53	0.18
With only 50 labeled samples			
S-CNN	15.63±5.45	12.25±4.22	16.29±6.24
S-CNN+PL	1.93±0.64	3.21±0.79	3.17±0.56
S-CNN+PL+TC^{avg}	1.29±0.81	3.08±0.44	1.08±0.29
S-CNN+PL+TC (**Ours**)	**0.36±0.28**	**0.69±0.39**	**0.64±0.27**

Methods of Face Anti Spoofing

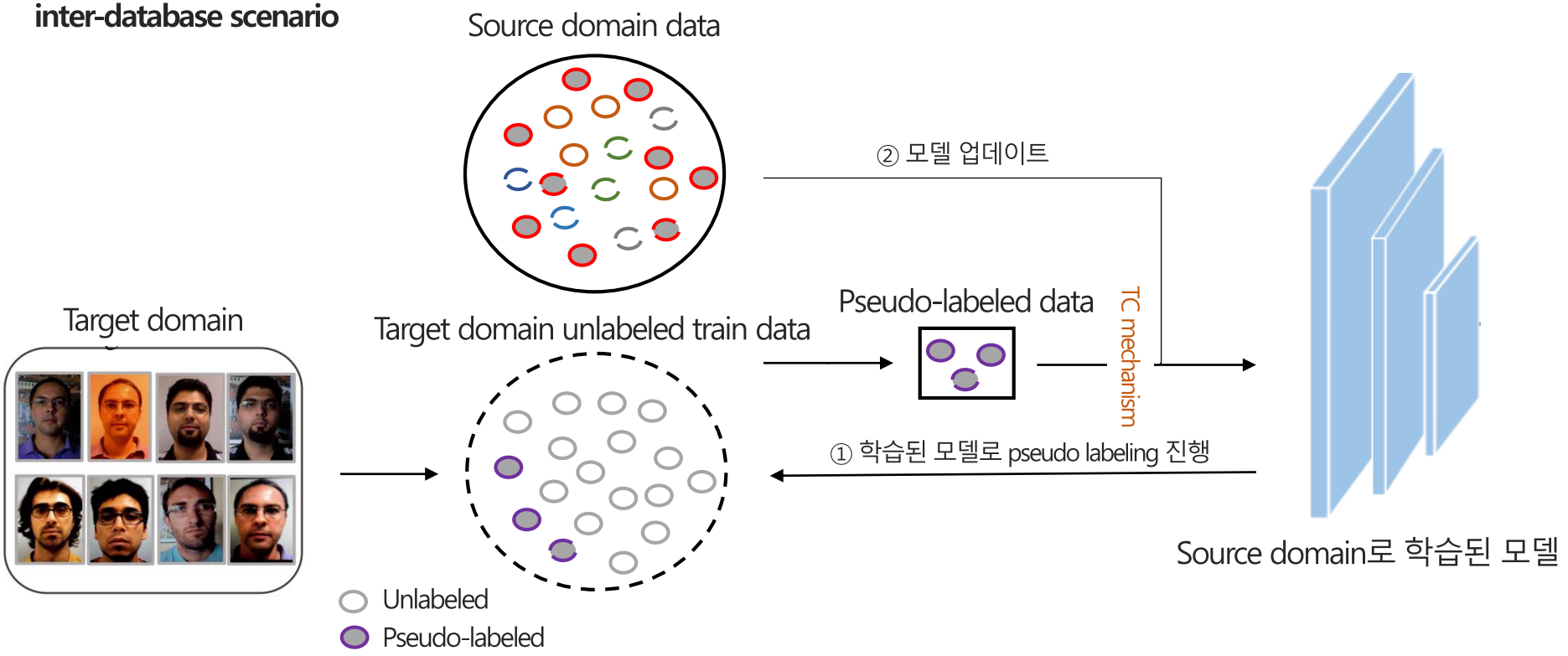
Progressive Transfer Learning for Face Anti-Spoofing

❖ 제안 방법론 – Adaptive Transfer mechanism

① Progressive learning (PL): 다수의 레이블 데이터를 얻기 어려운 문제를 해결하기 위한 실시간 학습 방법

② Adaptive Transfer (AT): 새로운 유형의 공격에 신속하게 대응하기 위한 방법

inter-database scenario



Methods of Face Anti Spoofing

Progressive Transfer Learning for Face Anti-Spoofing

❖ Experiment – inter database

- 본 연구에서 사용한 4개의 데이터 셋 중 3개를 소스 도메인으로 1개를 타겟 도메인으로 설정하여 실험 진행
- 소스 도메인 내에 모든 레이블 데이터를 다 사용하였을 때에도 가장 우수한 성능을 보임
- 소스 도메인마다 50개의 레이블 데이터만을 사용하여도 모든 레이블 데이터를 다 사용한 지도 학습 접근 방식 모델보다 우수한 성능을 보임

Method	O&C&I to M		O&M&I to C		O&C&M to I		I&C&M to O	
	HTER (%) ↓	AUC (%) ↑	HTER (%) ↓	AUC (%) ↑	HTER (%) ↓	AUC (%) ↑	HTER (%) ↓	AUC (%) ↑
All data from source domains are labeled								
Color Texture [39]	28.09	78.47	30.58	76.89	40.40	62.78	63.59	32.71
MADDG [33]	17.69	88.06	24.50	84.51	22.19	84.99	27.89	80.02
SSDG-R [42]	7.38	97.17	10.44	95.94	11.71	96.59	15.61	91.54
SSDG-R*	15.81±2.50	90.44±1.94	11.14±1.24	95.66±0.94	23.61±1.54	79.30±3.07	21.82±1.25	85.95±1.12
S-CNN	16.52±3.10	91.21±2.77	14.69±2.95	94.10±2.03	28.75±1.20	78.87±4.12	20.48±4.47	87.47±1.01
S-CNN+PL	12.83±3.65	92.07±3.19	10.11±3.04	95.18±2.66	19.55±3.28	88.82±4.52	18.04±3.58	88.91±2.13
S-CNN+PL+TC	12.31±2.73	94.63±2.85	7.85±2.24	97.34±1.99	16.97±2.45	91.67±4.29	16.61±3.10	90.56±2.76
S-CNN+PL+AT	9.67±1.55	95.81±1.25	7.32±1.62	97.78±0.59	14.27±1.03	95.21±2.19	16.18±1.40	92.45±1.74
S-CNN+PL+TC+AT (Ours)	**7.82±1.21**	**97.67±1.09**	**4.01±0.81**	**98.96±0.77**	**10.36±1.86**	**97.16±1.04**	**14.23±0.98**	**93.66±0.75**
Only 50 faces from each source domain are labeled								
S-CNN	21.70±5.88	86.74±6.36	37.71±7.12	68.44±8.45	29.32±4.32	72.46±5.18	32.18±5.23	73.93±6.52
S-CNN+PL	14.98±5.02	91.54±4.71	15.75±4.87	93.76±4.21	23.84±5.79	80.02±4.52	19.98±5.01	86.09±5.58
S-CNN+PL+TC	14.12±4.15	91.65±4.28	12.75±3.76	94.01±4.42	19.65±4.24	83.35±2.86	19.74±3.55	88.34±3.89
S-CNN+PL+AT	11.89±3.52	94.23±3.15	8.63±4.29	94.93±3.23	15.54±3.10	87.67±3.13	16.01±2.36	89.97±1.79
S-CNN+PL+TC+AT (Ours)	**10.73±3.76**	**96.56±2.67**	**6.51±3.22**	**96.12±3.34**	**14.89±2.39**	**93.11±2.15**	**15.73±1.97**	**91.96±1.43**

Methods of Face Anti Spoofing

Generalize to unseen attack type – domain generalization

❖ Domain Invariant Vision Transformer Learning for Face Anti-spoofing (2023, WACV)

- National Taiwan University에서 연구하였으며, 1월 5일 기준 11회 인용
- 해당 연구의 기여점은 크게 2가지로 요약될 수 있음
  - ① Domain invariant Vison Transformer (DiVT)를 제안함
  - ② 도메인에 상관없이 실제 얼굴(live)에 대한 특징 벡터를 특징 공간의 원점에 위치하도록 하는 손실 함수(L_{DiC}) 제안
  - ③ 도메인에 상관없이 공격 유형 별로 특징 공간 상에서 군집화하는 손실 함수(L_{DiA}^{ce}) 제안

Domain Invariant Vision Transformer Learning for Face Anti-spoofing

Chen-Hao Liao[1], Wen-Cheng Chen[2], Hsuan-Tung Liu[3], Yi-Ren Yeh[4], Min-Chun Hu[5], Chu-Song Chen[1]
[1]National Taiwan University, [2]National Cheng Kung University, [3]E.SUN Financial Holding Co., Ltd.,
[4]National Kaohsiung Normal University, [5]National Tsing Hua University
r09922113@csie.ntu.edu.tw, jerrywiston@mislab.csie.ncku.edu.tw, ahare-18342@esunbank.com.tw,
yryeh@nknu.edu.tw, anitahu@cs.nthu.edu.tw, chusong@csie.ntu.edu.tw

Methods of Face Anti Spoofing

Domain Invariant Vision Transformer Learning for Face Anti-spoofing

❖ 연구배경 및 문제 상황

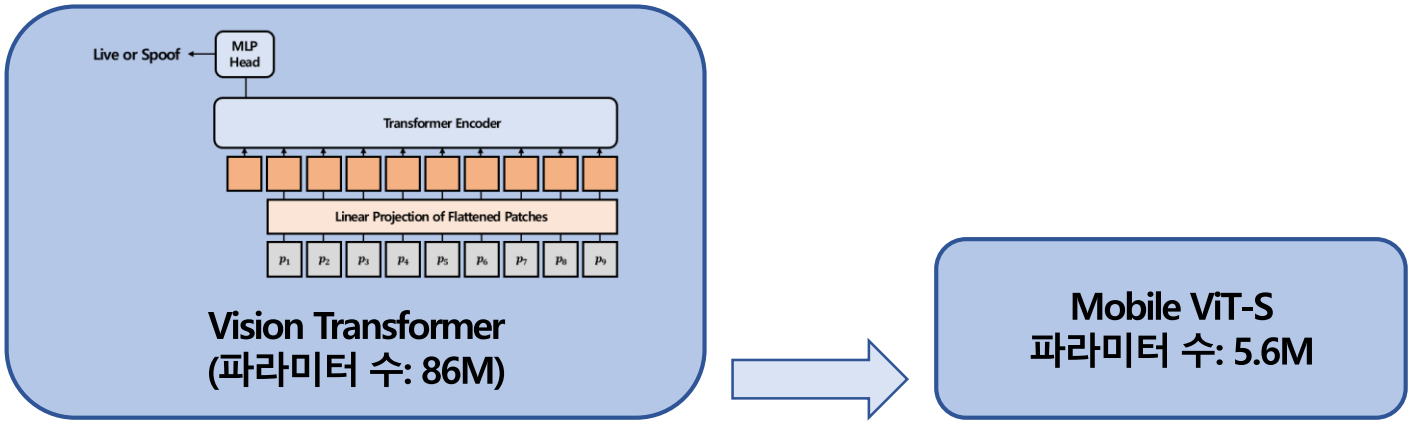
- 알려진 데이터 셋에서만 높은 성능을 달성하는 것이 아닌, 알려지지 않은 도메인의 데이터에 효과적인 일반화 필요
- 가짜 데이터(spoof)의 위조 패턴은 이미지 전체에 전역적으로 분포
- 이를 해결하기 위해,
  - ① 지역적인 정보에 집중하는 CNN이 아닌 전역적인 관점에서 이미지를 분석할 수 있는 사전 학습된 Mobile-ViT 활용
  - ② 도메인에 상관없이 실제 얼굴(live)에 대한 특징 벡터를 특징 공간의 원점에 위치하도록 하는 손실 함수(L_{DiC}) 제안
  - ③ 도메인에 상관없이 공격 유형 별로 특징 공간 상에서 군집화하는 손실 함수(L_{DiA}^{ce}) 제안

Methods of Face Anti Spoofing

Domain Invariant Vision Transformer Learning for Face Anti-spoofing

❖ 연구배경 및 문제 상황

- 알려진 데이터 셋에서만 높은 성능을 달성하는 것이 아닌, 알려지지 않은 도메인의 데이터에 효과적인 일반화 필요
- 가짜 데이터(spoof)의 위조 패턴은 이미지 전체에 전역적으로 분포
- 이를 해결하기 위해,
  - ① 지역적인 정보에 집중하는 CNN이 아닌 전역적인 관점에서 이미지를 분석할 수 있는 사전 학습된 Mobile ViT 활용
  - ② 도메인에 상관없이 실제 얼굴(live)에 대한 특징 벡터를 특징 공간의 원점에 위치하도록 하는 손실 함수(L_{DiC}) 제안
  - ③ 도메인에 상관없이 공격 유형 별로 특징 공간 상에서 군집화하는 손실 함수(L_{DiA}^{ce}) 제안



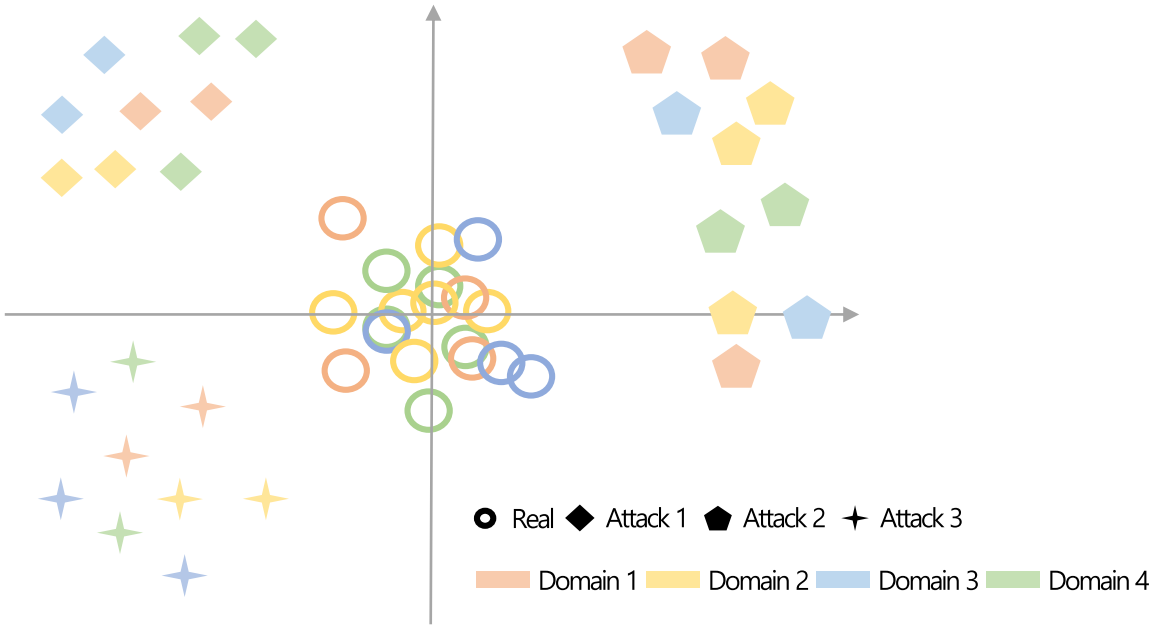
- Transformer 계열 모델을 사용했을 때의 2가지 한계점으로 인해 Mobile ViT 사용
  - ① Large model size
  - ② High computational cost

Methods of Face Anti Spoofing

Domain Invariant Vision Transformer Learning for Face Anti-spoofing

❖ 연구배경 및 문제 상황

- 알려진 데이터 셋에서만 높은 성능을 달성하는 것이 아닌, 알려지지 않은 도메인의 데이터에 효과적인 일반화 필요
- 가짜 데이터(spoof)의 위조 패턴은 이미지 전체에 전역적으로 분포
- 이를 해결하기 위해,
  - ① 지역적인 정보에 집중하는 CNN이 아닌 전역적인 관점에서 이미지를 분석할 수 있는 사전 학습된 Mobile ViT 활용
  - ② 도메인에 상관없이 실제 얼굴(live)에 대한 특징 벡터를 특징 공간의 원점에 위치하도록 하는 손실 함수(L_{DiC}) 제안
  - ③ 도메인에 상관없이 공격 유형 별로 특징 공간 상에서 군집화하는 손실 함수(L_{DiA}^{ce}) 제안



Methods of Face Anti Spoofing

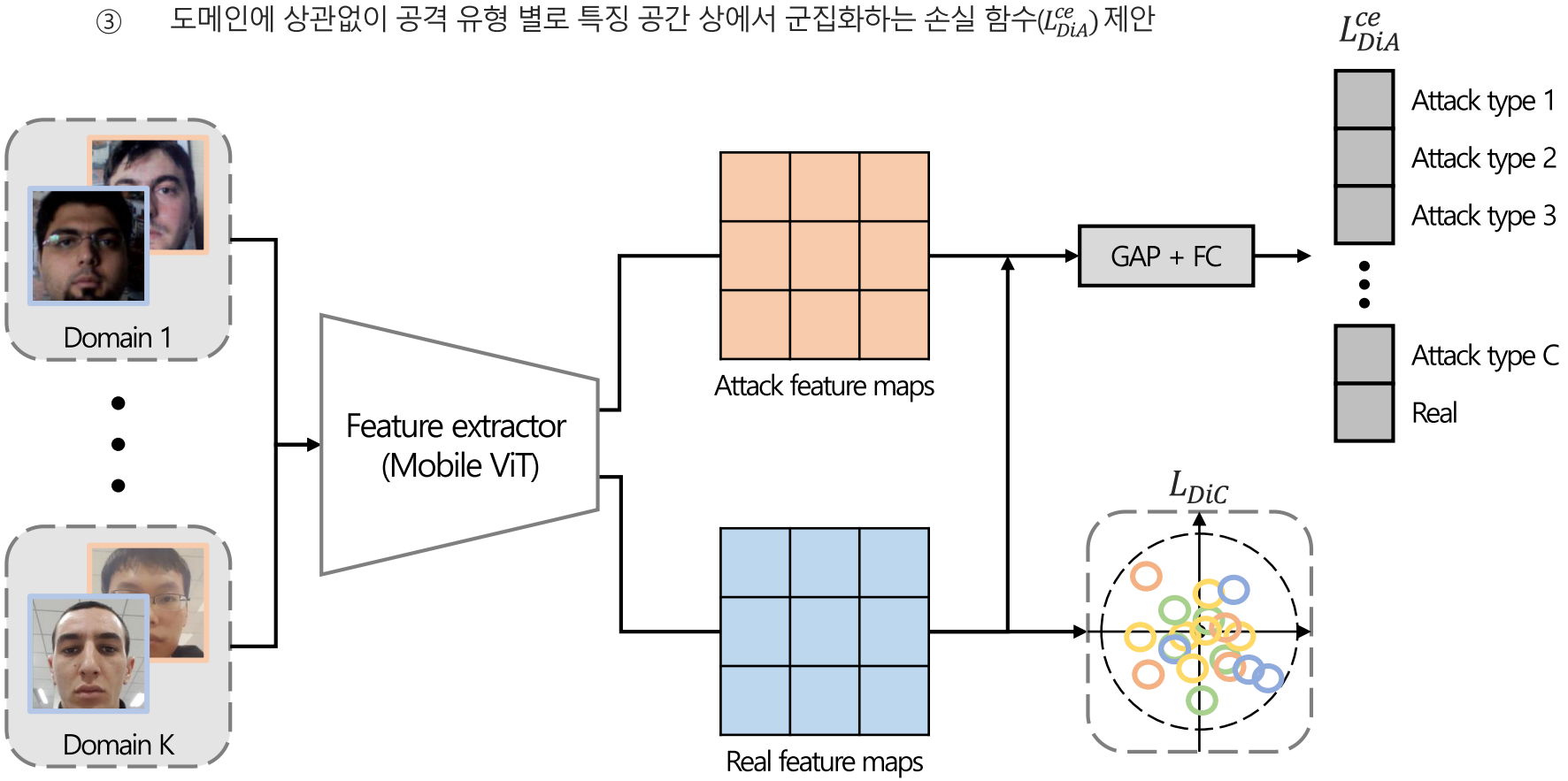
Domain Invariant Vision Transformer Learning for Face Anti-spoofing

❖ 제안 방법론 – L_{DiC} & L_{DiA}^{ce}

① 지역적인 정보에 집중하는 CNN이 아닌 전역적인 관점에서 이미지를 분석할 수 있는 사전 학습된 Mobile ViT 활용

② 도메인에 상관없이 실제 얼굴(live)에 대한 특징 벡터를 특징 공간의 원점에 위치하도록 하는 손실 함수(L_{DiC}) 제안

③ 도메인에 상관없이 공격 유형 별로 특징 공간 상에서 군집화하는 손실 함수(L_{DiA}^{ce}) 제안



Methods of Face Anti Spoofing

Domain Invariant Vision Transformer Learning for Face Anti-spoofing

❖ 제안 방법론 – L_{DiC} & L_{DiA}^{ce}

① 도메인에 상관없이 실제 얼굴(live)에 대한 특징 벡터를 특징 공간의 원점에 위치하도록 하는 손실 함수(L_{DiC}) 제안

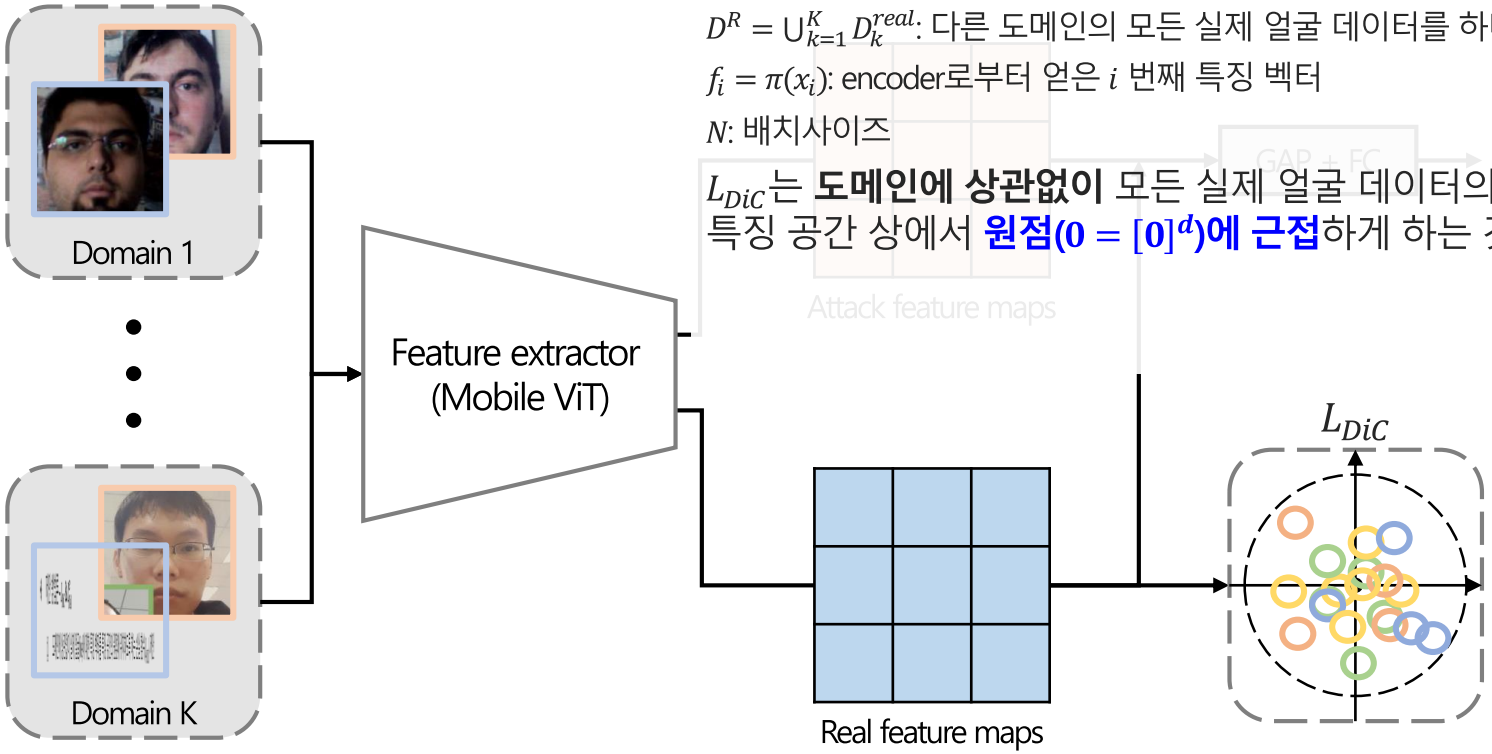
$$L_{DiC} = \frac{1}{N}\mathbb{I}[x_i \in D^R] \cdot \|f_i\|_1$$

$D^R = \cup_{k=1}^{K} D_k^{real}$: 다른 도메인의 모든 실제 얼굴 데이터를 하나로 결합함

$f_i = \pi(x_i)$: encoder로부터 얻은 i 번째 특징 벡터

N: 배치사이즈

L_{DiC}는 **도메인에 상관없이** 모든 실제 얼굴 데이터의 특징 벡터가 특징 공간 상에서 **원점($\mathbf{0} = [0]^d$)에 근접**하게 하는 것이 목표



Methods of Face Anti Spoofing

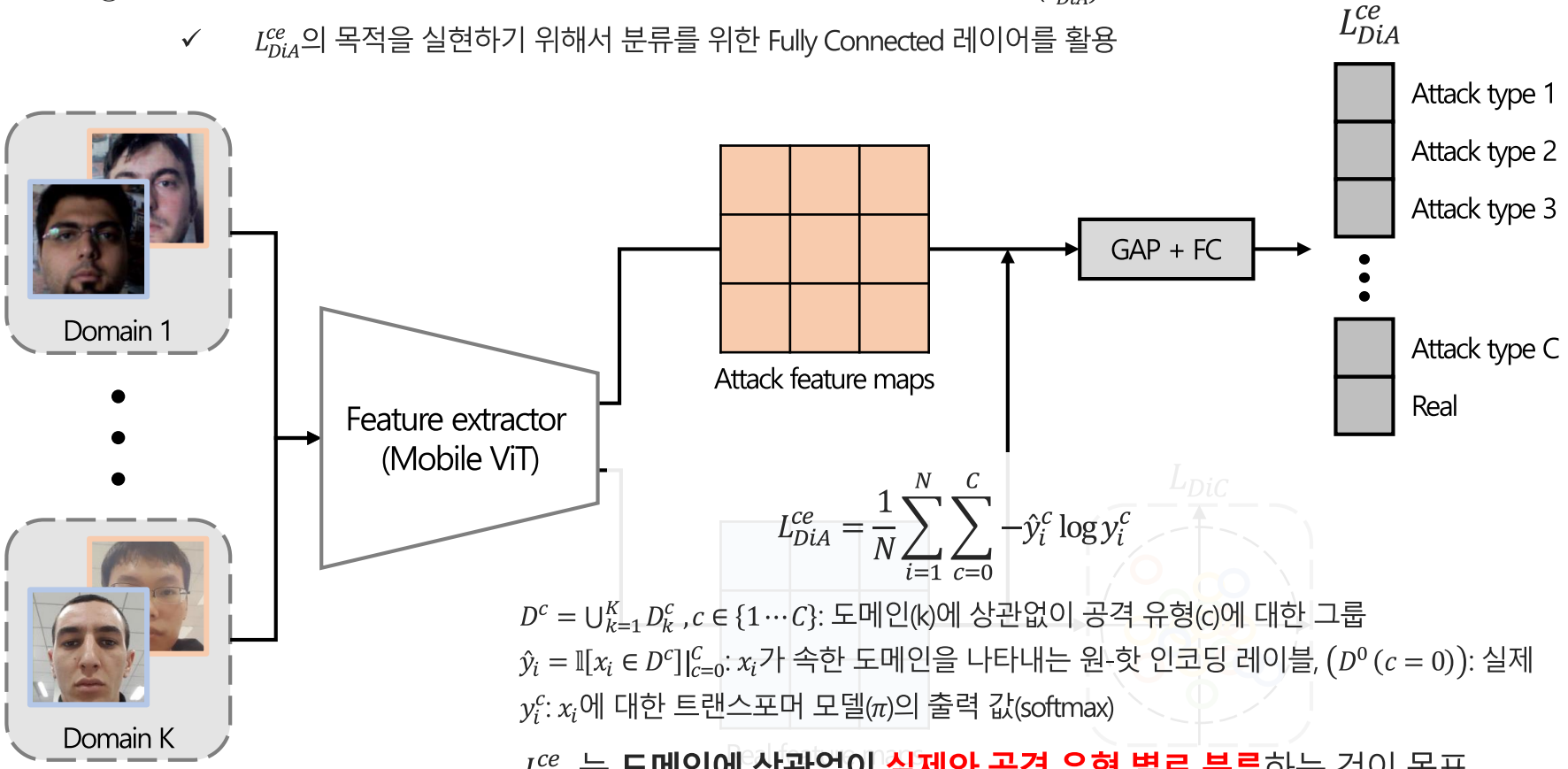
Domain Invariant Vision Transformer Learning for Face Anti-spoofing

❖ 제안 방법론 – L_{DiC} & L_{DiA}^{ce}

① 도메인에 상관없이 실제 얼굴(live)에 대한 특징 벡터를 특징 공간의 원점에 위치하도록 하는 손실 함수(L_{DiC}) 제안

② 도메인에 상관없이 공격 유형 별로 특징 공간 상에서 군집화하는 손실 함수(L_{DiA}^{ce}) 제안

✓ L_{DiA}^{ce}의 목적을 실현하기 위해서 분류를 위한 Fully Connected 레이어를 활용



$$L_{DiA}^{ce} = \frac{1}{N}\sum_{i=1}^{N}\sum_{c=0}^{C} -\hat{y}_i^c \log y_i^c$$

$D^c = \cup_{k=1}^{K} D_k^c, c \in \{1 \cdots C\}$: 도메인(k)에 상관없이 공격 유형(c)에 대한 그룹

$\hat{y}_i = \mathbb{I}[x_i \in D^c]|_{c=0}^{C}$: x_i가 속한 도메인을 나타내는 원-핫 인코딩 레이블, $(D^0\,(c=0))$: 실제

y_i^c: x_i에 대한 트랜스포머 모델(π)의 출력 값(softmax)

L_{DiA}^{ce}는 **도메인에 상관없이 실제와 공격 유형 별로 분류**하는 것이 목표

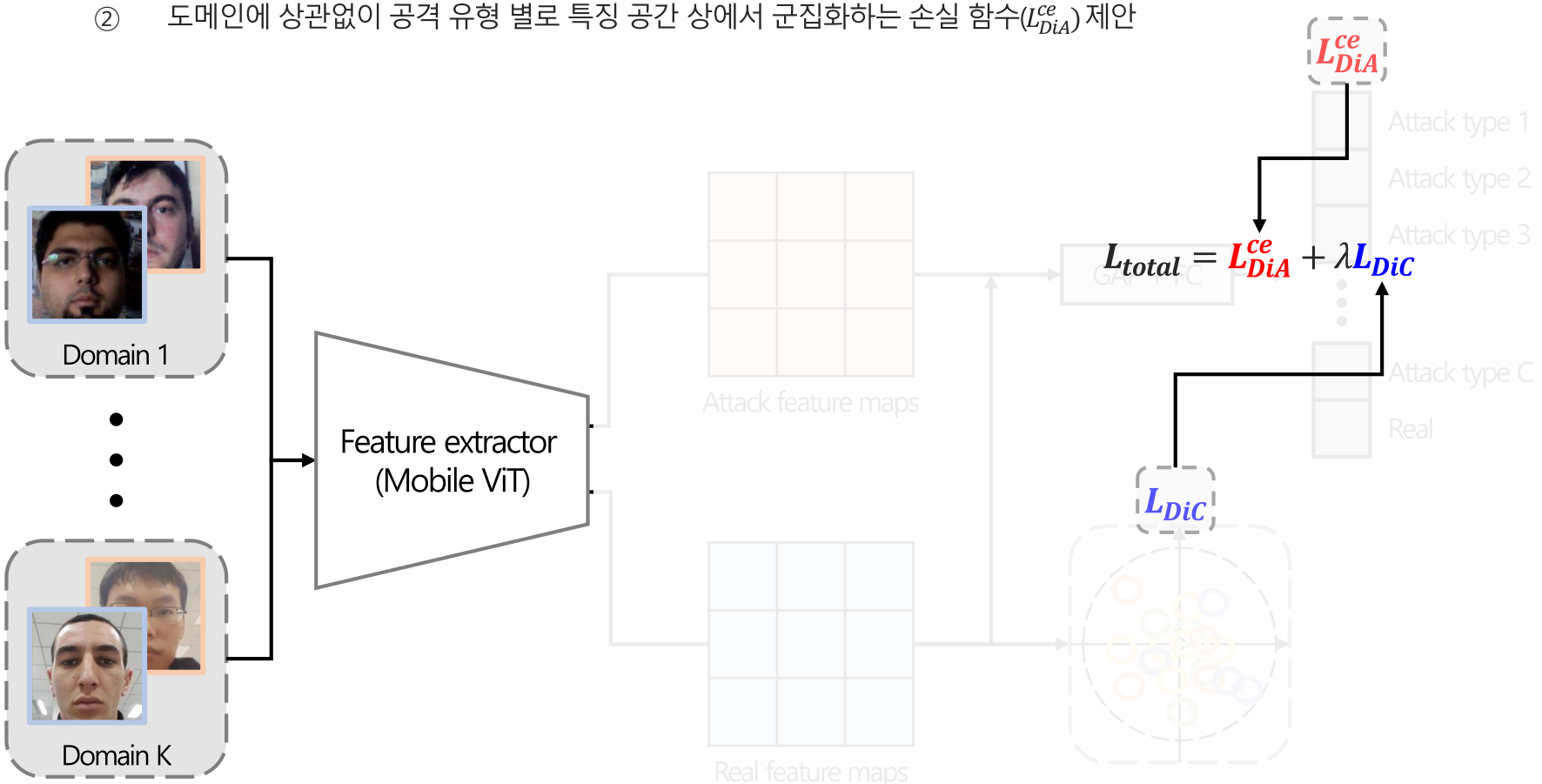
Methods of Face Anti Spoofing

Domain Invariant Vision Transformer Learning for Face Anti-spoofing

❖ 제안 방법론 – L_{DiC} & L_{DiA}^{ce}

① 도메인에 상관없이 실제 얼굴(live)에 대한 특징 벡터를 특징 공간의 원점에 위치하도록 하는 손실 함수(L_{DiC}) 제안

② 도메인에 상관없이 공격 유형 별로 특징 공간 상에서 군집화하는 손실 함수(L_{DiA}^{ce}) 제안



Methods of Face Anti Spoofing

Domain Invariant Vision Transformer Learning for Face Anti-spoofing

❖ Experiment 1

- 총 4개의 벤치마크 데이터 셋 중 3개는 학습 데이터, 1개는 평가 데이터로 활용했을 때 성능 비교 실험
- 앞선 선행 연구보다 제안 방법론의 우수한 성능을 확인

Methods	O&C&I to M		O&M&I to C		O&C&M to I		I&C&M to O	
	HTER(%)	AUC(%)	HTER(%)	AUC(%)	HTER(%)	AUC(%)	HTER(%)	AUC(%)
MADDG (CVPR' 19) [27]	17.69	88.06	24.50	84.51	22.19	84.99	27.98	80.02
DR-MD-Net (CVPR' 20) [30]	17.02	90.10	19.68	87.43	20.87	86.72	25.02	81.47
NAS-FAS (TPAMI' 20) [42]	16.85	90.42	15.21	92.64	11.63	96.98	13.16	**94.18**
RFMeta (AAAI' 20) [28]	13.89	93.98	20.27	88.16	17.30	90.48	16.45	91.16
D^2AM (AAAI' 21) [4]	12.70	95.66	20.98	85.58	15.43	91.22	15.27	90.87
DRDG (IJCAI' 21) [41]	12.43	95.81	19.05	88.79	15.56	91.79	15.63	91.75
ANRL (ACM MM' 21) [19]	10.83	96.75	17.85	89.26	16.03	91.04	15.67	91.90
FGHV (AAAI' 22) [18]	9.17	96.92	12.47	93.47	16.29	90.11	13.58	93.55
SSDG-R (CVPR' 20) [13]	7.38	97.17	10.44	95.94	11.71	96.59	15.61	91.54
SSAN-R (CVPR' 22) [32]	6.67	98.75	10.00	96.67	8.88	96.79	13.72	93.63
DiVT-M (Ours)	**2.86**	**99.14**	**8.67**	**96.92**	**3.71**	**99.29**	**13.06**	94.04

Methods of Face Anti Spoofing

Domain Invariant Vision Transformer Learning for Face Anti-spoofing

❖ Experiment 2

- 실험 1보다 학습 데이터의 규모를 줄여서 실험을 진행 (MSU-MFSD, Idiap Replay-Attack만 학습 데이터로 활용)
- 제안 방법론이 실험 1에서 가장 성능이 좋았던 선행 연구(2개)보다 우수한 성능을 보임
  - ✓ M&I to C 세팅에서는 HTER 기준으로 SSDG-R이 제안 방법론보다 성능이 좋음
  - ✓ 하지만, 제안 방법론이 AUC(거짓 음성과 거짓 양성 사이에 균형 정도)기준 0.25% 높기 때문에 일반적으로 더 좋다고 평가하고 있음

Methods	M&I to C		M&I to O	
	HTER(%)	AUC(%)	HTER(%)	AUC(%)
SSDG-R [13]	**19.86**	86.46	27.92	78.72
SSAN-R [32]	25.56	83.89	24.44	82.56
DiVT-M	20.11	**86.71**	**23.61**	**85.73**

Methods of Face Anti Spoofing

Domain Invariant Vision Transformer Learning for Face Anti-spoofing

❖ Experiment 3

- 본 연구에서 제안한 두 가지 손실 함수의 사용 유무에 따른 성능 비교 실험
- 두 가지 손실 함수를 모두 활용하였을 때, 모든 세팅에서 우수한 성능을 보이는 것을 확인

Components		O&C&I to M		O&M&I to C		O&C&M to I		I&C&M to O	
L_{DiA}^{ce}	L_{DiC}	HTER (%)	AUC (%)	HTER (%)	AUC (%)	HTER (%)	AUC (%)	HTER (%)	AUC (%)
		5.48	93.99	13.22	93.32	17.14	90.98	15.28	90.78
✓		**2.62**	99.10	9.33	96.40	7.71	96.92	15.42	91.52
	✓	5.71	98.36	10.00	96.80	17.86	88.88	13.33	**94.11**
✓	✓	2.86	**99.14**	**8.67**	**96.92**	**3.71**	**99.29**	**13.06**	94.04

Conclusion

❖ 결론

- Introduction to Face Anti Spoofing
  - 보안 및 인증 시스템에 악의적인 행위를 방지하기 위해 Face Anti-Spoofing에 대한 연구가 진행
  - 보지 못했던 공격 유형 및 도메인에 대한 일반화를 위한 연구가 활발히 진행

- Methods of Face Anti Spoofing

① Progressive Transfer Learning for Face Anti-Spoofing (unseen domain – domain adaptation)

- ✓ 적은 양의 레이블 데이터를 사용하는 semi-supervised learning 기반 framework 제안
- ✓ 보지 못했던 도메인 간 차이를 줄이기 위해 adaptive transfer mechanism 제안
- ✓ 강건하고 신뢰할 수 있는 pseudo label을 선정하기 위해 temporal confidence consistency mechanism을 제안

② Domain Invariant Vision Transformer Learning for Face Anti-spoofing (unseen attack types – domain generalization)

- ✓ Domain invariant Vison Transformer (DiVT)를 제안함
- ✓ 도메인에 상관없이 실제 얼굴(live)에 대한 특징 벡터를 특징 공간의 원점에 위치하도록 하는 손실 함수(L_{DiC}) 제안
- ✓ 도메인에 상관없이 공격 유형 별로 특징 공간 상에서 군집화하는 손실 함수(L_{DiA}^{ce}) 제안