Journal of Korea Multimedia Society Vol. 25, No. 10, October 2022(pp. 1375-1385)
https://doi.org/10.9717/kmms.2022.25.10.1375

단일 훈련 샘플만을 활용하는 준-지도학습 심층 도메인 적응 기반 얼굴인식 기술 개발

김경태[†], 최재영 [††]

Development of Semi-Supervised Deep Domain Adaptation Based Face Recognition Using Only a Single Training Sample

Kyeong Tae Kim[†], Jae Young Choi[††]

ABSTRACT

In this paper, we propose a semi-supervised domain adaptation solution to deal with practical face recognition (FR) scenarios where a single face image for each target identity (to be recognized) is only available in the training phase. Main goal of the proposed method is to reduce the discrepancy between the target and the source domain face images, which ultimately improves FR performances. The proposed method is based on the Domain Adatation network (DAN) using an MMD loss function to reduce the discrepancy between domains. In order to train more effectively, we develop a novel loss function learning strategy in which MMD loss and cross-entropy loss functions are adopted by using different weights according to the progress of each epoch during the learning. The proposed weight adoptation focuses on the training of the source domain in the initial learning phase to learn facial feature information such as eyes, nose, and mouth. After the initial learning is completed, the resulting feature information is used to training a deep network using the target domain images. To evaluate the effectiveness of the proposed method, FR performances were evaluated with pretrained model trained only with CASIA-webface (source images) and fine-tuned model trained only with FERET's gallery (target images) under the same FR scenarios. The experimental results showed that the proposed semi-supervised domain adaptation can be improved by 24.78% compared to the pre-trained model and 28.42% compared to the fine-tuned model. In addition, the proposed method outperformed other state-of-the-arts domain adaptation approaches by 9.41%.

Key words: Domain Adaptation, Face Recognition, Semi-Supervised Learning, Domain Discrepancy, Deep Neural Network

※ Corresponding Author : Jae Young Choi, Address: (17035) 81 Oedae-ro, Mohyeon-myeon, Cheoin-gu, Yongin-si, Gyeonggi-do, Korea, TEL : +82-31-330-4906, FAX : +82-31-330-4906, E-mail : jychoi@hufs.ac.kr
Receipt date : Jun. 27, 2022, Revision date : Aug. 25, 2022
Approval date : Sep. 14, 2022
[†] Division of Computer Engineering, Hankuk University of Foreign Studies (E-mail : rudxo9251@gmail.ac.kr)
[††] Division of Computer Engineering, Hankuk University of Foreign Studies

※ This research was supported by Hankuk University of Foreign Studies Research Fund.
※ This research was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. 2021R1A2C1092322).
※ This research was supported by the "Development of Wave Overtopping quantitative observation technology" funded by the Korea Institute of Marine Science & Technology Promotion(KIMST) (No. 20220180).

1. 서 론

인공지능 기술을 활용하는 얼굴인식은 신원 인증을 위한 가장 효율적인 생체 인식 기술이며, 군사, 금융, 공공 보안은 물론 일상생활과 같은 방대한 영역에서 널리 사용되고 있다. 그러나 벤치마크 데이터셋에서 완벽하게 수행되는 인공신경망은 실제 응용분야에서 얼굴인식 성능이 저하되는 일반화 오류가 발생할 수 있다. 연구문헌 [1]에 따르면, 유명인들의 260만장 이상의 얼굴로 구성된 이미지로 훈련된 VGG 모델[2]의 교차 벤츠마크 평가를 하였을 때 LFW[3]에서 98.95%, YTF[4]에서 97.30%를 달성하였다. 그러나 GBU[5] 얼굴인식 데이터 집합에서 Ugly, Bad 및 Good 파티션에 대해 26%, 52% 및 85%의 성능으로 매우 저조한 성능을 보였다. 주요 원인은 훈련 데이터(소스 데이터)와 테스트 데이터(타켓 데이터)간의 분포가 다르기 때문이다. 이를 도메인 이동이라고 한다. 이 도메인 이동의 예시는 Fig. 1에 나타내었다. Fig. 1(a)에서 CASIA-WebFace[6]의 이미지는 제약이 없는 환경에서 인터넷에서 수집되며 대부분의 인물은 주변 조명에서 촬영된 유명인 및 대중이다. Fig. 1(b)에서 FERET[7]은 제약이 있는 실험실 환경에서 수집되었으며 얼굴의 포즈가 변한다. 이 두 가지 데이터셋은 한눈에 보기에도 다른 조명, 다른 화질, 다른 포즈로 촬영되어있으며 다른 환경의 데이터셋이라는 것을 쉽게 눈으로 구분할 수 있다. 이는 상당한 불일치가 있음을 나타낸다.

도메인 이동으로 인한 문제를 해결하기 위해 가장 많이 사용되는 접근 방식은 라벨이 지정된 데이터로 사전 훈련된 심층 신경망의 매개 변수를 별도의 라벨을 지정한 테스트 데이터로 조정하는 것이다. 하지만 이러한 접근 방식은 많은 양의 라벨이 지정된 데이터를 필요하기 때문에 비용이 많이 들거나 실행 불가능하다. 요즘 최신 연구로서 비지도 도메인 적응(Unsupervised Domain Adaptation)[8]과 준-지도 도메인 적응(Semi-supervised Domain Adaptation)[9,10]은 이 문제를 해결하는 것을 목표로 한다[31]. 이때 준-지도 도메인 적응은 비지도 도메인 적응 문제의 변형으로, 오직 소스(훈련 도메인)의 라벨이 있는 샘플만 사용하는 비지도 도메인 적응과 달리 준-지도 도메인 적응은 소스(훈련) 도메인의 라벨이 있는 샘플과 타켓(테스트) 도메인의 일부 샘플을 사용하여 타켓(테스트) 도메인에 대한 예측 모델을 학습한다. 타켓 도메인에 라벨이 지정된 샘플 수가 적기 때문에 준-지도 도메인 적응은 비지도 도메인 적응에 비해 성능을 크게 향상시킬 수 있다.

얼굴 인식은 일반적인 객체 인식(예: CIFAR10 [11])에 대한 비지도 도메인 적응와 다르게 타켓 도메인에 대한 데이터를 추가적으로 수집하는 것은 현실적인 제약이 많이 수반되며 샘플 수집이 되더라도 모든 데이터에 대해 라벨을 일일이 달 수 없다. 그리고 소스 도메인과 타켓 도메인의 얼굴 클래스(ID)가 겹치지 않기 때문에 일반적인 비지도 도메인 적응 방법을 직접적으로 적용하여 해결하기가 힘들다. 이러한 제약조건 때문에 비지도 도메인 적응을 얼굴인식에 적용하기 어렵지만, 준-지도 도메인 적응으로



(a) (b)

Fig. 1. Some sample face images of used face databases (DBs). (a) CASIA-webface DB [6] collected in real life conditions without restrictions and (b) FERET DB [7] collected in a restricted laboratory environment.

는 타켓 도메인의 일부 데이터를 훈련에 사용할 수 있기 때문에 소스 도메인와 타켓 도메인의 얼굴 클래스(ID)가 겹치지 않는 문제를 해결할 수 있다. 특히 범죄자 조회, 신원 인증 등과 같은 얼굴인식 응용에서는 얼굴인식 엔진 학습에 활용되는 훈련 영상들과 실제 인식에 활용되는 테스트 영상들의 취득 기기와 영상 품질이 매우 상이하기 때문에 신뢰성 있는 얼굴인식 성능을 얻기 위해서는 준-지도 도메인 적응 기술을 필수적으로 활용해야 한다. 얼굴인식과 관련된 준-지도 도메인 적응 기술[10]은 다양한 시나리오에 적용이 가능하지만 신원확인과 같은 얼굴인식 시나리오[9,12]에 특히 적합하다. 신원 조회 시 ID 카드, 전자여권, 머그샷(mugshot)의 사진은 정면 얼굴이며 저해요소가 적기 때문에 좋은 품질을 가지고 있고 라벨이 이미 존재하지만 영상 개수가 매우 제한적이다. 반면 보안용 CCTV 및 인터넷에 경우 많은 영상들을 취득할 수 있지만 좋지 않은 품질을 가지고 있으며 신원 라벨(identity label)이 부재한 경우가 대부분이다. 따라서 본 논문에서는 신원 라벨이 존재하는 품질 좋은 얼굴 영상 샘플들의 개수는 매우 적고 동시에 신원 라벨이 부재하면서 품질이 나쁜 영상들이 대다수인 얼굴인식 응용 시나리오에서 얼굴 인식 성능을 개선하기 위한 방법을 제안한다. 제안된 방법은 라벨이 있는 대용량 데이터를 소스 도메인으로 사용하며, 타켓 도메인에서 클래스 당 한 장의 라벨이 있는 영상과 라벨이 없는 여러 장의 얼굴 영상으로 학습하는 얼굴인식을 위한 준-지도 도메인 적응 학습 방법을 제안한다. 제안된 방법은 다음과 같은 독창성을 가진다.

• 도메인의 분포를 일치시켜주는 MMD 손실함수와 클래스 분류를 위한 Cross-entropy 손실함수를 최적으로 결합하여 훈련하였다. 손실함수 결합 시 단순한 결합 방식이 아닌 훈련이 진행됨에 따라 각각의 손실함수에 최적 가중치를 주어 훈련을 진행한다.

• 준-지도 도메인 적응 훈련 시 타켓 도메인에서 클래스당 한 장의 라벨이 있는 영상만을 사용하였다. 실생활에서 타켓 도메인에 대해 클래스를 여러 장씩 라벨을 매기는 것은 높은 비용, 시간 및 인력이 필요하기 때문에 현실적으로 많은 제약 사항이 존재한다. 제안된 방법의 얼굴인식 시나리오는 클래스당 여러 장의 라벨을 사용하는 것이 아닌 단 한 장의 라벨이 있는 품질이 좋은 영상만을 사용한 경우이다.

제안된 방법은 공개된 대용량 데이터셋인 CASIA-webface[6]와 FERET[7]를 사용하여 실험을 수행하였다. CASIA-webface와 FERET은 서로 인식 대상들의 클래스가 겹치지 않기 때문에 CASIA-webface를 소스 도메인으로 FERET을 타켓 도메인으로 활용하였다. 또한 라벨이 있는 품질이 좋은 영상을 활용하기 위해 FERET의 갤러리 영상을 사용하였다. 제안된 방법의 공정한 비교를 위해 동일한 시나리오 환경에서 CASIA-webface으로만 훈련된 사전훈련(pretrained) 모델과 FERET의 갤러리 만으로 훈련된 미세조정(fine-tuned) 모델과 성능 비교 평가를 진행하였다. 실험결과 제안한 준-지도 도메인 적응 방법을 활용할 경우 사전훈련 모델과 비교했을 때 24.78%로 향상 되었으며, 미세조정된 모델과 비교했을 때 28.42% 향상되었다. 또한 최신 도메인 적응 기술과 비교하였을 때 9.41% 인식성능을 향상시킬 수 있음을 검증하였다.

2. 관련 연구

2.1 심층 도메인 적응(Deep Domain Adaptation) 기술

도메인 적응 기술은 분할(Segementation)[13], 텍스트 분류[14], 이미지 분류[15-17]와 같은 실제 응용 프로그램에 대한 도메인 불일치를 처리하는데 상당히 강력하다. 도메인 적응의 핵심은 두 개의 서로 다른 도메인의 분포 불일치를 줄이기 위해 특징 표현 또는 분류기를 조정한다. 최근 많은 연구를 통해 도메인 적응 모델을 설계하여 심층 아키텍처를 도메인 적응 시나리오에 적용하였다[12,18]. 연구문헌 [19]에서는 심층 합성곱 신경망을 도메인 적응 시나리오로 일반화하는 DAN 아키텍처를 제안하였다. DAN은 계층(Layer)의 특징 표현을 다른 도메인 분포가 일치될 수 있도록 MMD(Maximum Mean Discrepancy) 손실함수를 설계하였다. MMD는 각 도메인 간의 특징표현 분포의 차이를 측정하는 방법으로 이 차이를 최소화하여 도메인 불일치를 줄인다. [20]에서는 여러 도메인별 여러 계층의 특징 분포를 공동(Joint)으로 정렬하는 전이(Transfer) 신경망 JAN (Joint Adaptation Networks)을 제안하였다. DAN의 경우 각 계층에 대해 비교를 하였다면 JAN은 모든 계층의 특징표현 분포를 공동(Joint)으로 고려하여 공동 분포 공간에 정렬하는 JMMD를 설계하였다.

[21]에서는 인공신경망이 두 도메인간의 구분을 힘들게 하여 도메인간의 불일치를 줄이는 DANN (Domain Adversarial Neural Net)을 제안하였다. 신경망의 앞부분은 특징 추출을 위해 설계되었고 이후 2개의 완전 연결 계층(Fully-connected layer)에 서로 공유된다. 분류를 위한 완전 연결 계층 중 하나는 클래스 분류를 위한 계층이며, 다른 하나의 계층은 도메인을 분류하기 위한계층이다. 이때 신경망이 도메인의 분류를 옳지 못하게 하여 도메인간의 불일치를 줄인다. 핵심 아이디어는 도메인 이동을 명시적으로 해결하여 심층 구조의 최상위 계층에서 특징 이동성을 향상시키는 것이다. 따라서 달성된 피드포워드(feed forward) 신경망은 소스 및 타켓 도메인간의 분포 발산에 의해 방해받지 않고 타켓 도메인에 유효할 수 있다.

2.2 얼굴인식을 위한 심층 도메인 적응 기술

대부분의 얼굴인식 심층 도메인 적응 기술은 라벨이 없는 타켓 데이터를 라벨이 있는 소스 데이터를 통해 학습하고자 한다. SSPP-DAN[22]은 소스 도메인과 타켓 도메인의 신원(클래스)은 동일하지만 다른 얼굴인식 심층 도메인 적응 연구와 다르게 소스 도메인의 품질 좋은 영상이 한 장이라고 가정하고, 타켓 도메인의 영상은 품질이 안 좋은 영상이 다수라고 가정한다. SSPP-DAN은 소스 도메인의 품질 좋은 영상 한 장을 3차원 모델링을 통해 다양한 각도의 얼굴 영상으로 재생성하여 도메인-적대적 훈련(domain-adversarial training)을 진행하였다. [23]에서는 소스 도메인과 타켓 도메인의 신원이 동일할 때의 비지도 도메인 적응 얼굴인식 방법을 제시하였다. 라벨이 부여된 데이터가 부족하거나 없어 사용할 수 없는 타켓 도메인에 대해 소스 도메인의 정보를 활용하여 학습하였다. 이때 MMD 손실함수를 이용하여 타켓 도메인과 소스 도메인간의 불일치를 줄인다. 본 논문의 시나리오는 SSPP-DAN[22]과 동일하게 얼굴 영상 한 장을 타켓 도메인으로 하였지만, 소스 도메인과 타켓 도메인간의 신원이 동일하지 않기 때문에 [22,23]와 다른 시나리오다. [24]은 전이 학습을 통한 지도 도메인 적응 얼굴 인식 방법을 제안하였다. 소스 도메인에 대해 신경망을 훈련시킨 뒤에 특징 추출 신경망 부분만 사용하여 타켓 도메인에 대해 재학습을 하였다. 이때 Triplet 손실함수를 활용하여 같은 신원의 특징들은 서로 가깝게 만들어주며, 다른 신원의 특징끼리는 멀게 만든다. 하지만 제안한 방법은 소스 도메인으로 훈련된 신경망이 타켓 도메인으로 훈련되다보니 도메인간의 불일치가 생기게 되며 이는 일반화 성능이 하락하게 된다. CDA[25]는 소스 도메인과 타켓 도메인의 신원이 다를 때 군집화 기반 비지도 도메인 적응 얼굴 인식방법을 제시하였다. 우선 소스 도메인과 타켓 도메인의 신원이 다르더라도 데이터들의 분포를 일치시키기 위해 MMD 손실함수를 사용하였다. 이후 소스 도메인과 타켓 도메인의 불일치를 해결한 뒤 타켓 도메인의 특징을 추출하여 유사도를 비교한다. 유사도가 높은 특징들에 대해 가짜 라벨(Pseudo label)을 부여하였고 재학습을 하여 성능을 향상하였다. 얼굴인식 응용(예: 범죄자 조회, 신원 인증 등)에는 신뢰성 있는 얼굴인식이 중요하지만 가짜 라벨이 잘못 부여되어 얼굴인식 신뢰성이 저하될 수 있다.

본 연구에서는 이전 심층 도메인 적용 얼굴인식 연구들과는 차별적으로 소스 도메인과 타켓 도메인간의 신원(클래스)이 동일하지 않아도 심층 신경망 학습이 가능한 방법을 제시한다. 또한 가짜 라벨 생성에 오류가 발생해도 신뢰성 있는 얼굴인식을 확보하기 위한 심층 신경망 학습을 위해 타켓 데이터로부터 각 인식 대상 부류의 라벨링이 존재하는 오직 한 장의 영상을 학습으로 사용하는 준지도 도메인 적응 기술 방법을 제안한다. 그리고 도메인 적응이 효과적으로 이루어지도록 학습과정에서 결합 손실함수에 적절한 가중치를 부여하여 성능 향상을 이룰 수 있는 방법을 제시한다.

3. 제안 방법

Fig. 2는 제안한 방법의 전체 프로세스를 나타낸다. Fig. 2는 특히 MMD 손실함수와 Cross-entropy 손실함수를 결합한 학습 방법을 보여준다. Fig. 2에서 녹색 선은 소스 도메인에서 데이터 처리를 보여주며 파란 선과 적색 선은 타켓 도메인에서 데이터 처리를 의미한다. 이때 파란 선은 신원 라벨이 존재하는 타켓 도메인이고 적색 선은 신원 라벨이 부재하는 타켓 도메인이다. 백본(Backbone) 신경망은 VGG16 [2]과 ResNet50[26]를 활용하였으며, 신경망의 말단 계층인 완전 연결계층은 제거하고 소스 도메인에 관한 완전연결 계층과 타켓 도메인에 관한 완전 연결계

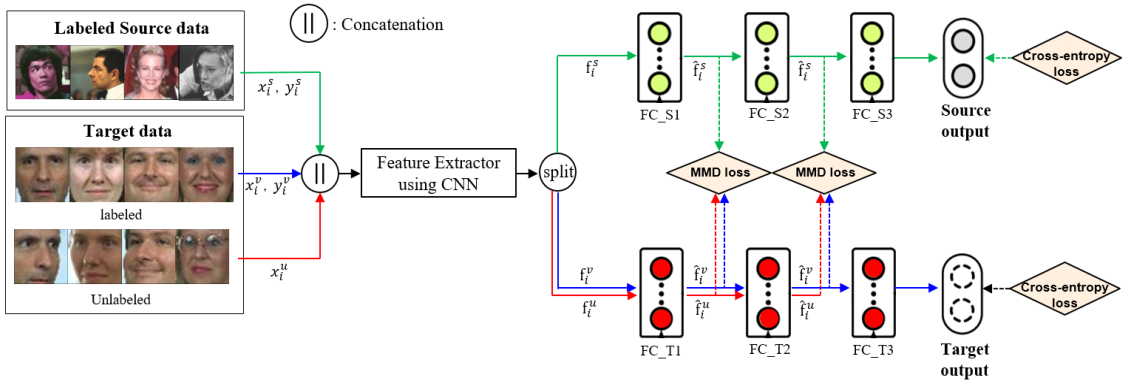


Fig. 2. Proposed pipeline process for implementing face recognition with deep domain adaptation using cross-entropy loss and MMD loss.

층을 연결하였다. 이후 각 완전연결 계층에서 나오는 특징들 간의 분포를 일치시키는 MMD 손실함수 2개와 각 말단의 완전연결 계층에 Cross-entropy 손실함수를 연결하였다. 테스트 방법은 학습된 신경망에서 소스 도메인에 관한 완전연결 계층은 제거하고 타켓 도메인에 관한 완전 연결 계층만 사용하여 타켓 도메인에 대한 신원 인식을 진행한다. 이후 갤러리(Gallery) 영상과 프로브(Probe) 영상이 입력으로 주어질 때 나오는 특징을 통해 유사도를 구하여 동일한 신원인지 여부를 판단한다.

3.1 준 지도 도메인 적응 신경망

준-지도 도메인 적응의 경우 전체 라벨이 지정된 소스 도메인 $D_s = \{(x_i^s, y_i^s)\}_{i=1}^{m_s}$과 일부 라벨이 있는 타켓 도메인이 있다. 타켓 도메인에서 라벨이 있는 샘플을 $D_v = \{(x_i^v, y_i^v)\}_{i=1}^{m_v}$이라 나타내며 라벨이 없는 샘플은 $D_u = \{(x_i^u)\}_{i=1}^{m_u}$으로 나타낸다. 이때 s는 소스, $t = \{v, u\}$는 타켓을 의미하며, m_s, m_v 그리고 m_u는 각 도메인의 샘플 수이다 일반적으로 샘플 수는 m_v보다 m_s, m_u가 크다. 제안한 시나리오에서는 라벨이 있는 타켓 데이터가 각 클래스마다 1장이 있다고 가정한다. 우리의 목표는 D_s, D_u 및 D_u의 훈련 세트에 대해 학습하고 D_u의 테스트 세트에 대해 평가하는 것이다.

제안된 방법은 크게 2가지 구조로 구성되어있다. 첫 번째는 특징 추출 파트와 도메인 별 특징 추출 파트로 구성되어 있다. 특징 추출 파트는 대용량으로 훈련된 신경망 모델을 신경망으로 VGG16[2] 모델과 ResNet50[26] 모델을 사용하였다. 해당 신경망은 대용량 데이터 셋으로 훈련되어 높은 성능을 보여준다. 제안방법에서 특징 추출 파트에서의 특징정보 추출은 다음과 같다.

$$\mathrm{f}_i^s, \mathrm{f}_i^v, \mathrm{f}_i^u = \theta_{\mathrm{FE}}(x_i^s, x_i^v, x_i^u)_{i=1, i=1, i=1}^{m_s, m_v, m_u} \tag{1}$$

여기서 θ_{FE}는 특징 추출을 위한 신경망이고, $\mathrm{f}_i^s, \mathrm{f}_j^v, \mathrm{f}_k^u$는 각각 전체 라벨이 있는 소스 도메인의 특징정보, 일부 라벨이 있는 타켓 도메인의 특징정보, 라벨이 없는 타켓 도메인의 특징정보를 나타낸다.

해당 특징정보들은 각 도메인별 특징 추출을 위한 개별적인 완전연결계층을 걸쳐 각 도메인 별 특징정보를 추출하게 된다.

$$\hat{\mathrm{f}}_i^s = \theta_{\mathrm{FC}}^{src}(\mathrm{f}_i^s)_{i=1}^{m_s} \tag{2}$$

$$\hat{\mathrm{f}}_i^v, \hat{\mathrm{f}}_i^u = \theta_{\mathrm{FC}}^{tgt}(\mathrm{f}_i^v, \mathrm{f}_i^u)_{i=1, i=1}^{m_v, m_u} \tag{3}$$

여기서 $\theta_{\mathrm{FC}}^{src}$는 소스 도메인의 특징 추출 및 분류를 위한 완전연결계층을 의미하며, $\theta_{\mathrm{FC}}^{tgt}$는 타켓 도메인의 특징 추출 및 분류를 위한 완전연결계층을 의미한다. 그리고 $\hat{\mathrm{f}}^s$는 소스 도메인의 특징 정보를 추출하기 위한 완전완결 계층의 특징 정보를 의미하며, $\hat{\mathrm{f}}^v, \hat{\mathrm{f}}^u$는 타켓 도메인의 특징 정보를 추출하기 위한 완전완결 계층의 특징 정보를 의미한다. 해당 특징정보는 MMD 손실함수[19]와 Cross-entropy 손실함수[27]를 사용하여 도메인 간의 격차를 줄이며 클래스를 분류하게 된다.

3.2 MMD(Maximum mean discrepancy) 손실함수

비지도 도메인 적응 분야에서 MMD[19]는 소스 도메인과 타겟 도메인 간의 불일치를 측정하기 위한 표준 분포 거리 메트릭으로 많이 사용되었다. 두 분포 s와 $t=\{v,u\}$가 주어지면 이들 사이의 MMD는 다음과 같이 정의된다.

$$L_M(s,t)=\sup_{\|\phi\|_{\mathcal{H}}\le 1}\| E_{x^s\sim s}[\Phi(\hat{\mathrm{f}}^s)]-E_{x^t\sim t}[\Phi(\hat{\mathrm{f}}^t)]\|^2_{\mathcal{H}} \quad (4)$$

여기서 E는 분포에 대한 기대치를 나타낸다. Φ는 원본 데이터를 재생 커널 힐베르트 공간(reproducing kernel Hilbert space, RKHS)에 매핑하는 함수를 나타낸다. MMD에서 정의한 통계를 기반으로 s와 t가 동일한 분포를 공유할 때 $MMD^2(s,t)=0$이다. 이 매핑과 연관된 커널 함수 $k(\hat{\mathrm{f}}^s,\hat{\mathrm{f}}^t)=\langle\Phi(\hat{\mathrm{f}}^s),\Phi(\hat{\mathrm{f}}^t)\rangle$는 n개의 PSD 커널 k_r의 볼록 조합 (Convex combination)으로 정의 된다.

$$K=\left\{k=\sum_{r=1}^{n}\beta_r k_r:\sum_{r=1}^{n}=1,\beta_r\ge 0,\forall r\right\} \quad (5)$$

여기서 β_r는 r번째 커널의 계수(coefficient)이고 일반적으로 사용되는 커널인 가우시안 커널 $k_r(\hat{\mathrm{f}}^s,\hat{\mathrm{f}}^t)=e^{-\|x_s-x_t\|^2/\gamma}$이다. $D_s=\{(x_i^s,y_i^s)\}_{i=1}^{m_s}$, $D_v=\{(x_i^v,y_i^v)\}_{i=1}^{m_v}$ 및 $D_u=\{(x_i^u)\}_{j=1}^{m_u}$라고 할 때 각각 분포 s와 $t=\{v,u\}$에서 도출되며, MMD의 경험적 추정은 다음과 같다.

$$L_M(D_s,D_t)=\left\|\frac{1}{m_s}\sum_{i=1}^{m_s}\Phi(\hat{\mathrm{f}}_i^s)-\frac{1}{m_v+m_u}\sum_{j=1}^{m_v+m_u}\Phi(\hat{\mathrm{f}}_j^t)\right\|^2_{\mathcal{H}} \quad (6)$$

MMD 기반 신경망, 즉 DDC[28] 및 DAN[19]의 주요 아이디어는 MMD 추정기(estimator)를 CNN 오류에 통합하여 도메인 발산을 최소화하는 것이다. 그러나 수식 (6)은 2차 시간 복잡도로 계산되며 깊은 신경망인 UDA의 경우 엄청나게 많은 시간이 소요된다. 참고문헌 [29]은 선형 복잡도를 갖는 MMD에 대한 편향되지 않은 근사를 추가로 제안했으며 미니 배치 방식의 기울기 계산에 적합하다. 제안된 MMD 손실함수는 다음과 같다.

$$\begin{aligned}L_M(D_s,D_t)=&\frac{1}{m_s(m_s-1)}\sum_{i\neq j}^{m_s}k(\hat{\mathrm{f}}_i^s,\hat{\mathrm{f}}_j^s)\\&+\frac{1}{(m_v+m_u)(m_v+m_u-1)}\sum_{i\neq j}^{m_v+m_u}k(\hat{\mathrm{f}}_i^t,\hat{\mathrm{f}}_j^t)\\&-\frac{2}{m_s(m_v+m_u)}\sum_{i,j=1}^{m_s,m_v+m_u}k(\hat{\mathrm{f}}_i^s,\hat{\mathrm{f}}_j^t)\end{aligned} \quad (7)$$

MMD에 의한 신경망 최적화를 통해 최종 분류 결정은 영역의 변화에 불변인 특성, 즉 소스 및 대상 도메인에서 동일하거나 매우 유사한 분포를 갖는 특성을 기반으로 이루어지므로 소스 데이터에 대해 학습된 모델은 타켓 데이터로 일반화할 수 있다.

3.3 Cross-entropy 손실함수

준-지도 도메인 적응에서 전체 라벨이 있는 소스 도메인과 일부 라벨이 있는 타켓 도메인의 클래스를 분류하기 위해 cross-entropy 손실함수를 사용한다. 도메인 적응 모델을 학습 할 때, 입력 데이터에 대한 소프트맥스 결과가 주어진 참 부류 레이블(ground-truth label)과 최대한 비슷해지도록 소프트맥스 cross-entropy 손실함수를 최소화 하는 방식이 주로 사용되는데, 각 도메인의 클라스를 분류하기 위해 두 가지 cross-entropy 손실 함수를 다음과 같이 정의한다.

$$\begin{aligned}L_s(D_s)=&-\sum_{i=1}^{N}\left(y_i^s\log(\hat{\mathrm{f}}_i^s))+(1-y_i^s)\log(1-(\hat{\mathrm{f}}_i^s)\right)\\L_v(D_v)=&-\sum_{i=1}^{C}\left(y_i^v\log(\hat{\mathrm{f}}_i^v))+(1-y_i^v)\log(1-(\hat{\mathrm{f}}_i^v)\right)\end{aligned} \quad (8)$$

여기서 i는 클래스의 인덱스이고, N은 소스 도메인에서 클래스의 총 개수를 나타낸다. 그리고 C는 타켓 도메인에서 클래스의 총 개수를 나타낸다.

위의 수식 (8)에서 나타낸 것 같이 각 손실 함수 L_s,L_v은 소스 도메인, 타켓 도메인의 참 부류 레이블에 대한 인식결과 오류(error)를 기준으로 한 cross-entropy 함수이다.

3.4 에포크 별 가중치가 결합 손실함수

본 연구에서는 준지도 학습을 효과적으로 진행하기 위해 에포크별 가중치를 다르게 설정하여 훈련하였다. 훈련 초기에는 소스 도메인에 대해 영상의 구체적인 구성요소(예: 눈, 코, 입 등)를 훈련시키기 위해 β와 δ는 0으로 설정하였다. 훈련 중기에는 타켓 도메인의 분류하는 완전연결 계층을 학습시키기 위해 α는 고정된 상태에서 β의 가중치를 높여 학습한다. 학습이 어느 정도 진행된 도메인 간의 분포를 줄이기 위해 α와 β는 고정된 상태에서 δ는 높여 학습이 된다. 제안한 방법의 최종 손실함수는 다음과 같다.

$$L_{total}=\alpha\times L_s+\beta\times L_v+\delta\times L_M \quad (9)$$

$$\alpha = 1$$

$$\beta = \begin{cases} 0 & \text{if } t \le t_0 \\ \left(\frac{t-t_0}{t_1-t_0}\right)^2 & \text{if } t_0 < t \le t_1 \\ 1 & \text{if } t_1 < t \end{cases}$$

$$\delta = \begin{cases} 0 & \text{if } t \le t_1 \\ \left(\frac{t-t_0}{t_2-t_0}\right)^2 & \text{if } t_1 < t \le t_2 \\ 1 & \text{if } t_2 < t \end{cases}$$

여기서 L_s는 소스 도메인에 대한 Cross-entropy 손실함수, L_t는 타켓 도메인에 대한 Cross-entropy 손실함수이며, L_M은 MMD 손실함수이다. 수식 (9)에서 제안한 가중치 결합손실함수를 활용해 소스 도메인인 대용량의 데이터로 훈련된 모델이 타켓 도메인의 부류들 사이에 특징 정보들을 잘 구분할 수 있도록 한다. 그리고 각 도메인에 특징 정보를 잘 구분한 상태에서 도메인간의 특징 분포 격차를 줄이며 도메인이 이동하기 때문에 성능향상을 기대할 수 있다. 만약 에포크에 따른 가중치 변화를 부여하지 않으면 훈련이 안된 신경망의 소스 도메인과 타켓 도메인 특징 정보는 일반적으로 비슷하여 같은 도메인 공간에서 소스 도메인과 타켓 도메인의 특징 정보를 구분하기 힘들어져 학습이 되지 않을 수 있다.

4. 실험 결과 및 고찰

이 절에서는 CASIA-webface[6], FERET[7] 데이터셋과 같은 2가지 얼굴 인식 데이터셋에서 제안방법을 평가한다. 본 연구에서 사용한 데이터 세트의 자세한 정보와 평가 프로토콜을 소개하는 것으로 시작하여 실험의 훈련 세부 사항을 설명하고 결과와 분석을 제시할 것이다.

4.1 실험 데이터셋과 평가 프로토콜

CASIA-WebFace: CASIA-WebFace 데이터셋[6]은 인터넷에서 수집한 대규모 얼굴 데이터 세트이다. 10,575개의 클래스와 494,414개의 이미지가 포함되어 있다. 라벨이 지정된 얼굴 데이터의 대규모는 CNN을 훈련하는 데 큰 도움이 됩니다. 본 실험에서는 이 데이터 세트를 분류 네트워크 훈련을 위한 소스 도메인 데이터로 채택하였다. CASIA-WebFace의 예제 샘플은 Fig. 1(a)에서 확인이 가능하다.

FERET: FERET 데이터셋[7]은 얼굴 인식을 위한 데이터 세트이다. 994개의 클래스와 11,338개의 이미지로 구성되어 있으며, 갤러리(Gallery)와 프로브(Probe)로 나눠져 있다. 본 실험에서는 70개의 클라스와 1,057장의 이미지를 무작위로 선택하여 실험을 진행하였다. 갤러리 영상은 정면을 응시한 영상들로 각 클래스당 1장씩 구성되어있고, 프로브 영상은 얼굴의 포즈, 조명이 다른 환경에서 촬영되었다. FERET의 예제 샘플은 Fig. 1(b)에서 확인이 가능하다.

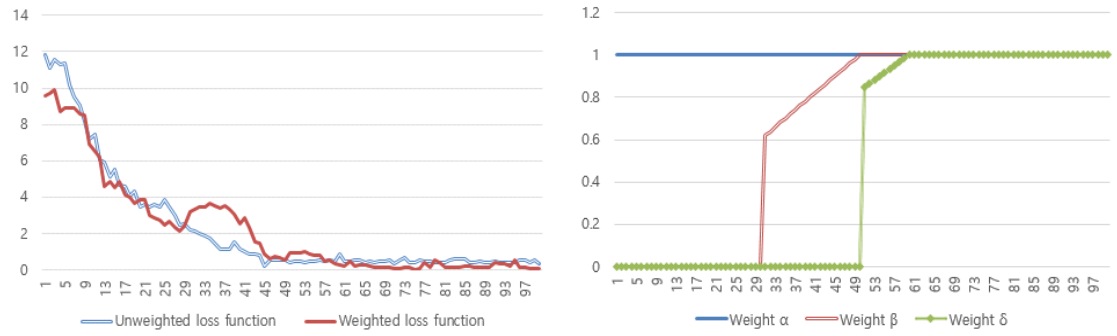
본 실험에서는 갤러리를 라벨이 있는 타켓 도메인 데이터로, 프로브 영상을 라벨이 없는 타켓 도메인 데이터로 채택하였다. FERET 데이터셋을 성능평가시 활용하였으며, FERET 데이터셋에서 갤러리와 프로브간의 유사도를 통해 Rank-1 성능을 도출하였다. 예를 들어 프로브 영상 중 하나와 갤러리 집합 내 모든 영상간의 유사도가 가장 높은 것을 선택하였고 선택된 영상이 갤러리와 동일한 라벨일 경우 Rank-1의 성능이 올라가게 된다.

4.2 실험환경

본 연구에서는 사전 훈련모델로 많이 사용되는 VGG16[2]와 ResNet50[26]를 사용하였으며 모든 입력 이미지는 224×224로 크기를 조정하였다. 그리고 배치 크기는 전체 라벨이 존재하는 소스 도메인은 36장, 라벨이 있는 타켓 도메인은 34장, 라벨이 없는 타켓 도메인은 2장으로 선택하였다. 또한 Adam optimizer[30]을 사용하였고 학습 속도(Learning rate)는 10^{-3}로 설정하였으며, 에포크(epoch)는 100으로 설정하였다. 수식 (9)의 에포크 별 가중치 적용을 위해 $t_0 = 30$, $t_1 = 50$, $t_2 = 60$으로 설정하였다. MMD의 커널은 가우시안 커널 $k(\hat{\mathrm{f}}^s, \hat{\mathrm{f}}^t) = e^{-\| x_s - x_t \|^2/\gamma}$이며, 여기서 γ은 대역폭(bandwidth)을 의미한다. 본 실험에서 각 계층 모두에 다중 커널 MMD를 적용한다. 5개의 가우시안 커널은 대역폭 γ_m을 $(1, 2^1, 2^2, 2^3, 2^4)$까지 설정하였으며, 여기서 γ_m은 훈련 데이터의 L-2 distance로 설정하였다.

4.3 가중치를 적용한 손실함수 실험결과

Fig. 3은 훈련이 진행됨(에포크)에 따라 손실함수의 변화를 나타내며 손실함수에 결합된 가중치가 얼마나 영향을 미치는지 보여주기 위한 그래프이다. Fig. 3(a)의 가중치가 적용되지 않은 손실 함수의 경



(a) (b)

Fig. 3. Change of proposed loss function value with respect to epoch. (a) training loss graph for each epoch with or without using weights and (b) weight of the function per epoch.

우 손실함수가 지속적으로 잘 떨어지며 수렴되지만, 특정 손실 값 약 0.4 이하로 떨어지지 않는다. 이는 훈련 시 대용량 공개 데이터로 훈련된 타켓 도메인에 대한 훈련이 충분히 이루어지지 않는 상태로 소스 도메인에 대한 훈련이 동시에 진행되기 때문에 얼굴의 특징적인 부분을 충분히 학습하지 않았기 때문이다. 반면 에포크별 가중치를 적용한 제안한 방법은 가중치를 적용하지 않는 방법보다 약 0.1이하로 손실값이 더 떨어진 것을 볼 수 있다. 이것은 소스 도메인에 대한 훈련과 타켓 도메인에 대한 훈련을 차례대로 진행되었고 충분히 학습이 되어 도메인 간의 이동이 더 적절하게 된 것이다. Fig. (b)는 Fig. (a)의 가중치가 적용된 함수로 훈련 시 각 에포크별 함수의 가중치 변화를 나타낸 그림이다.

4.4 FERET 데이터셋을 활용한 옵션 별 성능평가

제안된 방법은 타켓 도메인의 데이터셋인 FERET [7]으로 성능평가를 진행하였다. 공정한 성능 비교를 위해 사전 훈련 모델(Pretrained Model)과 미세조정된 모델(Fine-tuning Model)간의 비교를 진행하였다. 사전 훈련 모델은 백본 신경망에 소스 도메인만으로 학습 시킨 뒤 테스트 한 방법이며, 미세조정된 모델은 백본 신경망에 타켓 도메인의 갤러리 집합으로 학습 시킨 뒤 테스트한 방법이다. 또한 손실함수 간의 비교 평가를 위해 MMD 손실함수와 JMMD 손실함수를 각각 사용하여 실험하였다. 실험 결과 미세조정된 모델의 성능이 가장 저조하였다. 그 이유는 각 갤러리 집합에는 클래스당 한 장의 영상만으로 구성되어 있는데, 가장 품질이 좋은 영상이라도 적은 영상으로 훈련시키기에는 충분하지 않은 것으로 판

Table 1. Comparison of Rank-1 recognition performances with respect to different conditions on the target domain (FERET dataset).

Method	Backbone Network	Loss Function	Using weights	Target Rank-1 ACC.
Pretrained Model(S→T)	VGG16	CE loss	No	42.65%
Fine-tuning Model(T→T)		CE loss	No	37.89%
Proposed Method(S→T)		MMD[19]+CE loss	Yes	68.77%
Proposed Method(S→T)		JMMD[20]+CE loss	Yes	68.12%
Pretrained Model(S→T)	Resnet50	CE loss	No	44.72%
Fine-tuning Model(T→T)		CE loss	No	41.08%
Proposed Method(S→T)		MMD[19]+CE loss	Yes	69.50%
Proposed Method(S→T)		JMMD[20]+CE loss	Yes	69.09%

• S : source, T : target

단된다. 사전 모델의 성능의 경우 많은 데이터로 훈련되었기 때문에 미세 조정된 모델보다는 성능이 우수하나 제안한 방법보다는 성능이 떨어지는 것을 알 수 있다. 제안된 방법 중 백본 신경망으로 ResNet50 [26]를 활용하고 가중치가 적용된 결합 손실함수를 사용하였을 때 69.50%를 달성하였다. 이 성능은 사전 훈련된 모델보다 24.78% 성능 향상을 보였고, 미세 조정된 모델보다 28.42% 성능이 향상되었다. 이것은 소스 도메인인 대용량의 데이터로 훈련된 모델이 타켓 도메인에 특징 구분에 도움을 주어 타겟 도메인에 특징정보를 잘 구분하게 된다. 또한 각 도메인에 특징 정보를 잘 구분한 상태에서 특징정보를 이동하기 때문에 성능이 향상되었으며, 클래스 당 영상 한 장만으로 훈련이 충분히 가능하다는 것을 보인다.

4.5 최신 방법들과의 성능 비교

본 실험에서는 제안한 시나리오에 대한 방법의 우수성을 평가하기 위하여 최신 방법들과의 인식 성능을 비교하였다. Table 2는 최신 방법들과의 성능 비교를 나타내었다. 최신 딥 도메인 적응 방법으로서 DDC[28]는 도메인 이동을 해결하기 위해 최상위 계층에서 CNN 아키텍처와 단일 계층 적응을 활용한다. DAN[19]은 다중 커널 MMD와 여러 작업(task) 별 계층을 결합하여 더 많은 전이 가능한 특징을 학습한다. JMMD[20]는 다중 커널 MMD을 사용하며 계층을 함께 결합하여 전이 가능한 특징을 학습한다. 최신 방법들의 공정한 평가를 위해 동일한 소스 도메인과 타켓 도메인의 데이터를 사용하였고 준-지도학습 방법으로 진행하였다. 제안한 방법과 달리 DDC와 DAN과 JMMD는 가중치가 없는 손실함수로 실험을 진행하였다. 실험결과 제안한 시나리오에서 소스 도메인에 대한 성능은 대부분 97%이상으로 우수한 성능을 달성하였지만 타켓 도메인에 대한 성능은 DDC[28]가 59.14%, DAN[19] 가 58.85%, JMMD[20]가 60.09%로, 제안 도메인 적응 방법이 대략 9% 높은 얼굴인식 성능을 보였다. 이것은 클래스당 한 장의 라벨이 있는 영상만을 사용하는 준-지도 도메인 적응 방법에서 제안한 에포크 별 가중치가 성능향상에 큰 기여를 함을 논증한다.

Table 2. Comparison of Rank-1 recognition performance between the proposed method and other state-of-the-art methods on the CASIA-webface (source domain) and FERET datasets(target domain).

Method	Source Rank-1 ACC	Target Rank-1 ACC
DDC[28]	97.57%	59.14%
DAN[19]	96.99%	58.85%
JMMD[20]	97.85%	60.09%
Proposed Method	98.45%	69.50%

4. 결 론

본 논문은 소스 도메인과 타켓 도메인간의 불일치를 줄이되, 타켓 도메인의 클래스 당 라벨이 한 개씩 있는 고품질 영상이 존재 할 때의 시나리오에 대한 준 지도 학습 방법을 제안한다. 제안된 방법은 MMD 손실함수와 Cross-entropy 손실함수를 훈련의 진행함에 따라 가중치를 다르게 주어 효과적으로 훈련이 되도록 하였다. 훈련에 진행함에 따라 가중치를 적용한 제안한 방법은 가중치를 적용하지 않는 방법보다 오인식 비율이 더 낮은 것을 볼 수 있다. 이것은 소스 도메인에 대한 훈련과 타켓 도메인에 대한 훈련을 차례대로 진행되었고 충분히 학습이 되어 도메인 간의 이동이 더 적절하게 된 것이다. 제안된 방법의 공정한 비교를 위해 동일한 시나리오 환경에서 CASIA-webface으로만 훈련된 사전훈련(pretrained) 모델과 FERET의 갤러리 만으로 훈련된 미세조정(fine-tuned) 모델과 성능 비교 평가를 진행하였다. 실험결과 제안한 준-지도 도메인 적응 방법은 사전훈련 모델과 비교했을 때 24.78%로 향상 되었으며, 미세조정된 모델과 비교했을 때 28.42% 향상되었다. 또한 최신 준-지도 도메인 적응 기술과 비교하였을 때 9.41% 인식 성능이 향상되었다. 추후 연구에서는 타켓 도메인의 가짜 라벨 생성을 통해 재학습을 하는 자가-지도 학습(self-supervised learning)을 적용할 예정이다.

REFERENCE

[1] P.J. Phillips, “A Cross Benchmark Assessment of a Deep Convolutional Neural Network for Face Recognition,” *Proceedings of the Twelfth IEEE International Conference on Automatic Face & Gesture Recognition*, pp.

705-710, 2017.

[2] K. Simonyan, and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *arXiv Preprint*, arXiv:1409.1556, 2014.

[3] G.B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled Faces in the Wild: Adatabase for Studying Face Recognition in Unconstrained Environments," *Technical Report*, 07-49, University of Massachusetts, Amherst, 2007.

[4] L. Wolf, T. Hassner, and I. Maoz, "Face Recognition in Unconstrained Videos with Matched Background Similarity," *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pp. 529-534, 2011.

[5] P.J. Phillips, J.R. Beveridge, B.A. Draper, G. Givens, A.J. O'Toole, and D. Bolme, "The Good, the Bad, and the Ugly Facechallenge Problem," *Image and Vision Computing*. Vol. 30, No. 3, pp. 177-185, 2012.

[6] D. Yi, Z. Lei, S. Liao, and S.Z. Li, "Learning Face Representation from Scratch," *arXiv Preprint*, arXiv:1411.7923, 2014.

[7] P.J. Phillips, H. Moon, S.A. Rizvi, and P.J. Rauss, "The FERET Evaluation Methodology for Face Recognition Algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 10, pp. 1090-1104, 2000.

[8] R. Li, Q. Jiao, W. Cao, H.S. Wong, and S. Wu, "Model Adaptation: Unsupervised Domain Adaptation without Source Data," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9641-9650, 2020.

[9] G. He, X. Liu, F. Fan, and J. You, "Classification-Aware Semi-Supervised Domain Adaptation," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 964-965, 2020.

[10] K. Saito, D. Kim, S. Sclaroff, T. Darrell, and K. Saenko, "Semi-Supervised Domain Adaptation via Minimax Entropy," *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 8050-8058, 2019.

[11] A. Krizhevsky and G.E. Hinton, *Learning Multiple Layers of Features from Tiny Images*, *Master Thesis, Department of Computer Science*, University of Toronto, 2009.

[12] K. Bousmalis, G. Trigeorgis, N. Silberman, D. Krishnan, and D. Erhan, "Domain Separation Networks," *Advances in Neural Information Processing Systems*, pp. 343-351, 2016.

[13] Y. Zou, Z. Yu, B.V.K. Kumar, and J. Wang, "Unsupervised Domain Adaptation for Semantic Segmentation via Class-Balanced Self-Training," *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 289-305, 2018.

[14] H. Guo, R. Pasunuru, and M. Bansal, "Multi-Source Domain Adaptation for Text Classification via Distancenet-Bandits," *Proceedings of the AAAI Conference on Artificial Intelligence,* Vol. 34, No. 5, pp. 7830-7838, 2020.

[15] G. Matasci, M. Volpi, M. Kanevski, L. Bruzzone, and D. Tuia, "Semisupervised Transfer Component Analysis for Domain Adaptation in Remote Sensing Image Classification," I*EEE Transactions on Geoscience and Remote Sensing*, Vol. 53, No. 7, pp. 3550-3564, 2015.

[16] C. Deng, X. Liu, C. Li, and D. Tao, "Active Multi-Kernel Domain Adaptation for Hyperspectral Image Classification," *Pattern Recognition*, Vol. 77, pp. 306-315, 2018.

[17] J. Peng, W. Sun, L. Ma, and Q. Du, "Discrimi-Native Transfer Joint Matching for Domain Adaptation in Hyperspectral Image Classification," *IEEE Geoscience and Remote Sensing Letters*, Vol. 16, No. 6, pp. 972-976, 2019.

[18] W. Wang, R. Arora, K. Livescu, and J. Bilmes, "On Deep Multi-View Representation Learning," *International Conference on Machine*

Learning, pp. 1083-1092, 2015.

[19] M. Long, Y. Cao, J. Wang, and M.I. Jordan, "Learning Transferable Features with Deep Adaptation Networks," *Proceedings of the International Conference on International Conference on Machine Learning*, pp. 97-105, 2015.

[20] M. Long, H. Zhu, J. Wang, and M.I. Jordan, "Deep Transfer Learning with Joint Adaptation Networks," *International Conference on Machine Learning,* pp. 2208-2217, 2017.

[21] Y. Ganin, E Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, and V. Lempitsky, "Domain-Adversarial Training of Neural Networks," *The Journal of Machine Learning Research*, Vol. 17, No. 1, pp. 2096-2030, 2016.

[22] S. Hong, W. Im, J. Ryu, and H.S. Yang, "SSPP-DAN: Deep Domain Adaptation Network for Face Recognition with Single Sample per Person," *IEEE International Conference on Image Processing (ICIP)*, pp. 825-829, 2017.

[23] Z. Luo, J. Hu, W. Deng, and H. Shen, "Deep Unsupervised Domain Adaptation for Face Recognition," *13th IEEE International Conference on Automatic Face & Gesture Recognition,* pp. 453-457, 2018.

[24] G. Wen, H. Chen, D. Cai, and X. He, "Improving Face Recognition with Domain Adaptation," *Neurocomputing*, Vol. 287, pp. 45-51, 2018.

[25] M. Wang and D. Weihong, "Deep Face Recognition with Clustering Based Domain Adaptation," *Neurocomputing*, Vol. 393, pp. 1-14, 2020.

[26] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," *Proceedings of The IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770-778, 2016.

[27] Z. Zhang and M. Sabuncu, "Generalized Cross Entropy Loss for Training Deep Neural Networks with Noisy Labels," *Advances in Neural Information Processing Systems* 31 *(NeurIPS 2018)*, 2018.

[28] E. Tzeng, J. Hoffman, N. Zhang, K. Saenko, and T. Darrell, "Deep Domain Confusion: Maximizing for Domain Invariance," *arXiv Preprint*, arXiv:1412.3474, 2014.

[29] A. Gretton, K.M. Borgwardt, M.J. Rasch, B. Schölkopf, and A. Smola, "A Kernel Two-Sample Test," *The Journal of Machine Learning Research*, Vol. 13, No. 1, pp. 723-773, 2012.

[30] D.P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *arXiv Preprint*, arXiv: 1412.6980, 2014.

[31] T.C. Kang and H.B. Kang, "Machine Learning-Based Estimation of the Concentration of Fine Particulate Matter Using Domain Adaptation Method," *Journal of Korea Multimedia Society*, Vol. 20, No. 8, pp. 1208-1215, 2017.



김 경 태

2016년 중원대학교 학사
2017년~2019년 한국외국어대학교 컴퓨터공학과 석사
2019년~현재 한국외국어대학교 컴퓨터공학과 박사과정
관심분야 : 머신러닝, 패턴인식, 영상처리



최 재 영

2011년 KAIST 전기 및 전자공학과 박사
2008년~2012년 토론토대학 연구원
2012년~2013년 펜실베니아대학 연구원
2013년~2014년 삼성전자 책임연구원
2014년~2016년 중원대학교 의료공학과 조교수
2016년~현재 한국외국어대학교 컴퓨터공학부 부교수
관심분야 : 딥러닝, 머신러닝, 패턴인식, 영상처리