Journal of The Korea Institute of Information Security & Cryptology ISSN 1598-3986(Print)
VOL.27, NO.2, Apr. 2017 ISSN 2288-2715(Online)
https://doi.org/10.13089/JKIISC.2017.27.2.305

위조지문 판별률 향상을 위한 학습데이터 혼합 증강 방법*

김 원 진,† 김 성 빈, 유 경 송, 김 학 일‡
인하대학교

Data Mixing Augmentation Method for Improving Fake Fingerprint Detection Rate*

Weonjin Kim,† Cheng-Bin Jin, Jinsong Liu, Hakil Kim‡
Inha University

요 약

최근 모바일 및 핀테크(fin-tech) 분야의 최신 트렌드로 지문인식, 홍채인식과 같은 생체인식을 통한 사용자 본인 인증이 주목 받고 있다. 특히 지문인식을 이용한 인증 방식은 전통적인 생체인식 방식으로써 사용자들이 사용하는데 발생하는 거부감이 다른 생체인식에 비해 현저히 낮아 현재 가장 보편적으로 이용되는 방식이다. 이와 동시에 지문을 이용한 인증 시 보안에 대한 중요성이 부각되어 지문의 위조 여부 판별의 중요성 또한 증가하고 있다. 본 논문에서는 CNN(Convolutional Neural Networks) 특징을 이용한 위조 여부 판별 방법에 있어 판별률을 향상시키기 위한 새로운 방법을 제시한다. 학습데이터에 영향을 많이 받는 CNN 특성 상 기존에는 판별률을 향상시키기 위해 아핀 변환(affine transformation) 또는 수평 반전(horizontal reflection)을 사용하여 학습데이터의 양을 증가 시키는 것이 일반적인 방법이었으나 본 논문에서는 위조지문 판별 난이도를 기반으로 한 효과적인 학습데이터 증강(data augmentation) 방법을 제시하며 실험을 통해 제안하는 방법의 타당성을 확인하였다.

ABSTRACT

Recently, user authentication through biometric traits such as fingerprint and iris raise more and more attention especially in mobile commerce and fin-tech fields. In particular, commercialized authentication methods using fingerprint recognition are widely utilized mainly because customers are more adopted and used to fingerprint recognition applications. In the meantime, the security issues caused by fingerprint falsification bring lots of attention. In this paper, we propose a new method to improve the performance of fake fingerprint detection using CNN(Convolutional Neural Network). It is common practice to increase the amount of learning data by using affine transformation or horizontal reflection to improve the detection rate in CNN characteristics that are influenced by learning data. However, in this paper we propose an effective data augmentation method based on the database difficulty level. The experimental results confirm the validity of proposed method.

Keywords: fake fingerprint detection, data augmentation, CNN

I. 서 론

생체인식 기술은 살아 있는 사람의 생리적 또는 행동에 대한 특징을 인식하는 방법이다[1]. 일반적으로 지문, 목소리, 홍채, 망막, 손, 얼굴, 친필 등 다양한 종류의 생체인식 특성들이 생체인식 시스템에 폭넓게 사용되고 있다. 그 중 지문 인식 방법은 정확도, 인식 속도, 견고성 세 가지 요소 모두 균형이 잘

Received(02. 20. 2017), Modified(04. 04. 2017), Accepted(04. 04. 2017)
* 이 논문은 2016년도 정부(미래창조과학부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임. (no.R0190-15-2007, 스마트 디바이스용 박막 타입 지문센서 모듈 및 프라이버시 보호 응용 SW 기술개발)
† 주저자, wjkim@inha.edu
‡ 교신저자, hikim@inha.ac.kr(Corresponding author)

맞는 가장 대중적인 생체인식 방법이다. 또한, 상대적으로 편의성이 떨어지는 공인인증서와 보안카드를 대신하여 간단히 지문을 이용한 인증 시스템이 모바일 Fin-tech 관련 어플리케이션들에 이용되고 있다.[2] 하지만 이와 같은 지문 인증 시스템은 젤라틴, 라텍스, 실리콘과 같이 주변에서 쉽게 구할 수 있는 물질을 이용하여 다른 사람의 지문을 어렵지 않게 위조할 수 있다는 취약점이 존재하며, 국내에서는 이와 관련한 사건·사고가 지속적으로 발생하고 있다. 따라서 지문 이미지에 대한 위조 여부를 판단할 수 있는 기술은 지문 인증 시스템에서 매우 중요한 기술이라고 할 수 있다[3]. 이에 최근에 지문 이미지의 위조 여부 판별에 독보적인 판별률을 보이고 있는 CNN을 이용한 위조지문 판별 방법에 있어 학습데이터 증강 관점에서 성능 향상을 위해 학습을 원하는 위조지문 데이터베이스보다 상대적으로 판별 난이도가 높은 데이터베이스를 혼합하는 방식인 위조지문 판별 난이도를 기반으로 한 새로운 방안을 제시한다.

본 논문의 구성은 다음과 같다. 2장에서는 위조지문 판별 및 데이터 증강 방법 관련 연구에 대해 설명하였으며, 3장에서는 본 논문에서 사용한 위조지문 판별 방법에 대한 간략한 소개와 본 논문의 핵심 제안 내용인 위조지문 판별률 향상을 위한 학습데이터 혼합(mixing) 증강 방법에 대하여 설명하였다. 4장에서는 앞서 제안한 방법에 대한 실험환경 및 실험방법, 실험결과를 확인하였다. 마지막으로 5장에서는 결론 및 향후 연구에 대해 기술하며 논문을 마친다. 본 논문의 의의는 다음과 같다.

첫째, 지문 이미지의 위조 여부 판별에 대한 성능 향상을 위해 기존 CNN을 활용한 이미지 데이터 분류에 있어 사용되었던 일반적인 학습데이터 증강 방법을 사용하는 것이 아닌, 위조지문 판별 난이도를 기반으로 한 새로운 방법을 제시하였다.

둘째, 제안하는 학습데이터 증강 방법과 기존의 방법과 비교해보았을 때 제안하는 방법이 더욱 효과적인 방법임을 실험을 통해 증명하였다.

II. 관련 연구

2.1 위조 지문 판별 방법

위조 지문 이미지의 위조 여부를 판별하는 방법으로는 크게 하드웨어 기반인 방법과 소프트웨어 기반인 방법이 있다. 먼저 하드웨어 기반의 방법은 혈압[4], 피부 변형[5], 채취[6]와 같은 특정한 특징들을 추출할 수 있는 추가적인 센서를 필수적으로 요구하므로 효용성이 떨어진다. 소프트웨어 기반 방법은 위조 여부를 판별하기 위해 추가적인 센서 없이 지문 센서로부터 입력받은 지문 이미지로부터 특징들을 추출한 뒤 이를 이용하여 위조 여부를 판별한다. 따라서 전반적으로 소프트웨어 기반 방법이 하드웨어 기반 방법보다 비교적으로 저렴하고 유연하다.

소프트웨어 기반 방법에서 특징을 이용하는 방법으로 Jain 등[7]은 융선(ridge)의 밀도 및 크기, 연속성과 같은 세부적인 특징들을 이용하여 위조 여부를 판별하였다. 반대로 세부적인 지문의 특징을 이용하는 것이 아닌 이미지로부터 일반적인 특징을 추출하는 방법들이 존재한다. 그러한 방법 중 하나로 지문 이미지의 밝기 값, 대조 값, 주파수 영역에서의 위상 값을 이용한 방법이 있다[8]. 또한, 이미지의 texture 정보를 추출 할 때 사용하는 Weber Local Descriptor를 지문 이미지에 적용하기도 하였다[9]. 다른 방법으로는 LBP(Local Binary Patterns)를 사용한 방법과 변형된 LBP를 사용한 방법을 이용하여 성공적으로 위조 여부를 판별하기도 하였다[10][11][12].

위에서 설명한 방법들은 연구자들의 관찰과 실험에 의해 설계된 특징들을 사용한 방법이다. 이와 달리 CNN(Convolutional Neural Network)과 같이 학습을 통해 데이터로부터 추출한 특징을 사용하는 방법이 존재하며, 이러한 방법은 현재 지문 이미지의 위조 여부 판별에 있어 가장 높은 판별률을 보이고 있다.

Nogueira 등[13]은 지문 이미지의 위조 여부 판별을 위해 처음으로 CNN을 사용하여 위조 지문 이미지 판별에 성공하였다. 비슷하게 Wang 등[14]은 CNN 특징을 기반으로 한 투표 전략(voting strategy)을 사용하여 위조 지문 이미지 판별 성능을 향상시켰다. Nogueira 등[15]은 VGGnet을 사용하여 최근에 개최 된 위조지문 판별 대회인 LivDet2015에서 훌륭한 성능을 보였다. 최근에 CNN을 사용한 위조지문 판별 관련 연구로는 Chang 등[16]이 CNN 및 LBP, LPQ 특성을 활용하여 스마트폰 지문 센서로부터 취득한 지문 이미지에 대해 위조 여부 판별에 성공한 연구가 있으며, Park 등[17]이 작은 사이즈의 지문 패치를 활용하여 매우 높은 판별률로 위조 여부를 판별하였다.

2.2 학습데이터 증강 방법

학습데이터 증강(data augmentation)이란 알고리즘 성능에 있어 학습데이터 크기에 영향을 많이 받는 CNN 특성상 원본 학습데이터의 크기를 인위적으로 늘리는 방법이다. 학습데이터 증강 방법을 사용한 후 분류기(classifier)를 학습시키면 과적합(overfitting) 문제를 해결해주며 분류기를 보다 보편적이고 중요한 특징들에 대해 학습하도록 만들어 견고하고 성능 변동의 폭이 작은 분류기를 학습 결과로써 얻을 수 있다. 이러한 학습데이터 증강 방법은 CNN을 활용한 분류(classification)에 있어 성능 향상을 위해 여러 연구에서 사용되었다[15][18][20][21][22].

Nogueira 등[14]은 학습데이터 증강 방법으로 원본 이미지데이터의 각 모서리와 중앙을 기준으로 80% 영역만큼을 잘라낸 뒤 수평 반전까지 주어 원본 데이터 크기보다 학습 데이터 크기를 다섯 배까지 늘려 학습에 사용하였다. Krizhevsky 등[18]은 ImageNet 이미지 데이터[19]들에 대한 분류 성능을 향상시키기 위해 수평 반전(horizontal reflection) 및 RGB 채널(channel) 값들에 변화를 주며 학습데이터를 증강했다. Ciregan 등[20]은 다양한 종류의 이미지 데이터 분류에 있어 처음으로 사람과 비교할만한 결과를 얻기 위한 방법 중 하나로 이미지 데이터들의 무작위 이동(random translation), 크기 변환, 회전(rotation) 등과 같은 증강 방법을 사용하여 학습데이터의 크기를 증가시켰다. Ciresan 등[21]과 Simard 등[22]는 분류기의 성능을 향상시키기 위한 학습데이터 증강 방법으로 아핀 변환(affine transformation)과 왜곡(deformation)을 활용하였다.

본 논문에서는 CNN을 활용한 위조 지문 이미지 판별에 관한 성능을 향상시키기 위해 기존에 주로 사용된 학습데이터 증강 방법보다 효율적인 새로운 방법을 제시한다. 이어 실험을 통해 제안하는 방법의 타당성을 확인하고 결과를 분석한다.

III. 제안하는 방법

3.1 CNN을 활용한 위조 여부 판별

본 논문에서 위조지문 이미지의 위조 여부 판별을 위해 사용한 CNN 모델은 Fig. 1.과 같이 VGG 네트워크 구조를 기본으로 하여 학습 속도 및 시각화 결과를 고려한 변형 모델을 사용하였다[23][24].

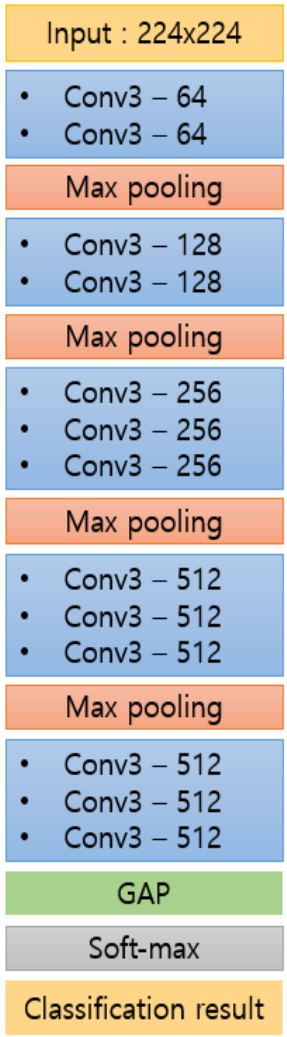

Fig. 1. Architecture of CNN

학습데이터 증강 방법에 대한 다양한 종류의 실험을 효율적으로 진행하기 위해서는 학습 속도가 중요하므로 학습 속도를 향상시키기 위해 우선 완전 연결계층(fully connected layer)을 모두 제거하였으며 GAP(Global Average Pooling)를 사용하였다.

모델 학습은 기본적인 지문 이미지에 대한 전처리 과정인 지문영역 추출과정을 거친 지문 이미지들을 입력 받아 진행 된다. Fig. 2.와 같이 모델 학습을 마친 뒤, 학습에 사용하지 않은 실제 지문 이미지와 위조 지문 이미지를 활용한 테스트 과정으로부터 위조 여부 판별이 성공적으로 됨을 확인 할 수 있다.

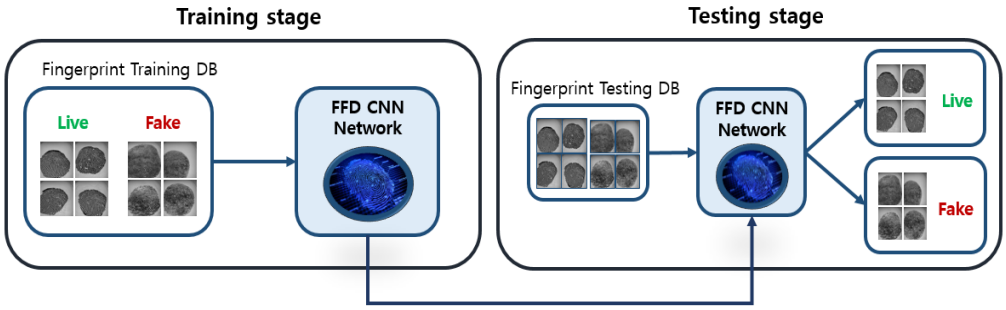

Fig. 2. Liveness detection using CNN

3.2 취득 센서에 따른 학습데이터 혼합 증강 방법

3.2.1 상이한 센서의 데이터 혼합

LivDet 데이터베이스는 위조지문 관련 공인데이터베이스로써 취득 센서에 따라 구분되어 있으며, 각 센서 별 dpi 및 지문 이미지의 크기, 위조지문 구성 물질이 다르다. 이를 이용하여 여러 종류의 위조지문 데이터베이스를 혼합한 새로운 학습데이터베이스를 구성한 뒤 CNN 모델을 학습시켜 위조지문 데이터베이스 혼합에 따른 위조지문 판별률의 변화추이를 분석하였다.

3.2.2 센서 특성 별 데이터 혼합

각 연도별 LivDet 데이터베이스의 구성을 살펴보면 지문 센서의 제조사가 같아 취득된 지문 이미지의 특성이 비슷한 데이터베이스들이 존재한다. 이를 활용하여 동일한 센서 제조사의 위조지문 데이터베이스를 혼합하여 학습데이터베이스를 구성한 뒤 CNN 모델을 학습시켜 위조지문 판별률과의 상관관계를 분석하였다.

3.3 판별 난이도에 따른 학습데이터 혼합 증강 방법

3.3.1 판별 난이도 기준 데이터 분류

본 논문에서 실험에 사용한 데이터베이스는 LivDet 2011, LivDet 2013, LivDet 2015이다.[25][26][27]. 각 데이터베이스를 구성하고 있는 하위 위조지문 데이터베이스들의 state-of-the-art ACE(Average Classification Error) 값을 Table 1.과 같은 기준으로 Table 2.와 같이 난이도 별로 분류하였다.

Table 1. Level of database difficulty

ACE(%)	Level
0.0 ~ 1.0	1
1.0 ~ 2.0	2
2.0 ~ 3.0	3
3.0 ~ 4.0	4
4.0 ~ 5.0	5
5.0 ~ 6.0	6
6.0 ~ 7.0	7
7.0 ~ 8.0	8
8.0 ~ 9.0	9
9.0 ~ 10.0	10

Table 2. LivDet databases with difficulty level

DB	Sensor	ACE(%)	Level
LivDet 2011	Biometrika	4.9 [8]	5
	Digital	2.0 [28]	3
	Italdata	8.0 [15]	9
	Sagem	1.7 [15]	2
LivDet 2013	Biometrika	0.8 [15]	1
	Crossmatch	3.2 [15]	4
	Italdata	0.4 [15]	1
	Swipe	2.8 [15]	3
LivDet 2015	CrossMatch	1.9 [15]	2
	Digital Persona	6.28 [15]	7
	GreenBit	4.8 [27]	5
	Hi Scan	4.2 [27]	5

3.3.2 판별 난이도에 따른 데이터 혼합

위조지문 데이터베이스들을 혼합하여 새로운 학습 데이터베이스를 구성 시 위조 여부 판별 난이도를 고려한다. CNN 모델의 구성 파라미터들을 위조지문 판별에 더욱 엄격하게 학습시키기 위해 위조 여부 판별 난이도가 낮은 위조지문 데이터베이스를 기준으로 삼아 상대적으로 난이도가 높은 위조지문 데이터베이스를 혼합시키며 학습데이터를 증강한 뒤 CNN 모델을 학습시킨다. 학습 후 테스트 과정을 통해 데이터 증강에 따른 판별률 변화를 확인하였다. 보편적인 이론 설립을 위해 반대로 위조 여부 판별 난이도가 높은 위조지문 데이터베이스를 기준으로 상대적으로 난이도가 낮은 위조지문 데이터베이스를 혼합시켜가며 데이터베이스 혼합에 따른 판별률 변화를 확인하였다.

IV. 실 험

4.1 실험 환경

실험은 CNN 모델 구현 용이성과 학습속도를 고려하여 구글에서 개발한 머신러닝을 위한 오픈소스 소프트웨어 라이브러리인 텐서플로우(tensorflow)를 사용하였으며 학습과 테스트 과정 모두에 그래픽 카드를 사용하였다. 학습 과정에서 사용한 그래픽카드로는 NVIDIA GTX 1080을 사용하였다. 실험에 사용한 데이터베이스 LivDet 2011, LivDet2013, LivDet 2015는 세부적으로 Table 2.와 같이 각기 다른 센서로부터 취득된 데이터베이스로 구성 되어 있으며 Biometrika, Crossmatch, Italdata 각 센서별 데이터들의 특성은 Table 3., Table 4., Table 5.와 같다.

Table 3. Property of Biometrika data

Sensor	DB	
Biometrika	LivDet 2011	LivDet 2013
dpi	500	569
Image size	312*372	317*372
Fake materials	5	5
Training DB size	2000	2000
Testing DB size	2000	2000

Table 4. Property of Crossmatch data

Sensor	DB	
Crossmatch	LivDet 2013	LivDet 2015
dpi	500	569
Image size	800*750	800*750
Fake materials	4	5
Training DB size	2250	2983
Testing DB size	2250	2351

Table 5. Property of Italdata data

Sensor	DB	
Italdata	LivDet 2011	LivDet 2013
dpi	500	500
Image size	640*480	640*480
Fake materials	5	5
Training DB size	2000	1990
Testing DB size	2000	2000

Table 6. Experimental results of Biomerika & Crossmatch data mixing

Training		Testing(ACE,%)	
		Bio	Cro
LivDet 2013	Bio	1.75	-
	Cro	-	4.37
	Bio+Cro	1.65	5.70

Table 7. Experimental results of Biometrika & Italdata data mixing

Training		Testing(ACE,%)	
		Bio	Ita
LivDet 2013	Bio	1.75	-
	Ita	-	1.80
	Bio+Ita	1.95	1.40

Table 8. Experimental results of Italdata & Crossmatch data mixing

Training		Testing(ACE,%)	
		Ita	Cro
LivDet 2013	Ita	1.80	-
	Cro	-	4.37
	Ita + Cro	1.70	5.73

4.2 실험 방법

3.1에서 언급한 CNN을 활용한 위조 여부 판별 알고리즘을 학습시키는데 3.2와 3.3에서 제안하는 학습데이터 혼합 증강 방법을 사용하였으며 학습 후 테스트 과정을 통해 지문의 위조 여부 판별률을 확인하였다. 또한 기존에 보편적으로 사용되었던 학습데이터 증강 방법과의 비교를 위해 기존 방법을 활용하여 데이터를 증강 시킨 비교용 학습데이터를 구성하였다. 이를 사용하여 CNN 모델을 학습시킨 성능과 제안하는 방법의 성능을 비교·분석하였다.

4.3 실험 결과

4.3.1 상이한 센서의 위조지문 데이터 혼합 실험 결과

서로 상이한 센서로부터 취득된 지문 이미지들로 구성된 위조지문 데이터베이스들을 혼합한 뒤 학습시켜 얻은 성능에 대한 결과는 Table 6, Table 7, Table 8 와 같다. 실험 결과 서로 다른 두 가지 센서로부터 취득된 지문 이미지로 구성된 데이터베이스를 혼합하여 학습시킨 다음 다시 각 센서 별 테스트 데이터베이스를 활용하여 위조 여부 판별 성능 평가를 할 경우 하나의 모델은 성능이 좋아지는 반면 나머지 하나는 성능이 오히려 떨어진다는 결과를 확인하였다.

4.3.2 센서 특성 별 데이터 혼합 실험 결과

LivDet 데이터베이스의 구성을 보면 연도별로 센서 제조사가 같아 구성하고 있는 이미지의 특성이 비슷한 데이터베이스들이 존재한다. 따라서 센서 특성에 따른 데이터 혼합 실험 결과를 비교하기 위해 Biometrika, Crossmatch, Italdat 센서로부터 취득한 이미지들로 구성 된 데이터베이스들을 선택하여 실험한 결과는 Table 9, Table 10, Table 11과 같다. Table 9 Biometrika 데이터 혼합 실험 결과를 보면 데이터베이스 혼합 후 LivDet 2011에 대한 성능은 감소하였으나, LivDet 2013에 대한 성능은 향상 된 것을 확인하였다. 이와 비슷한 결과를 Table 10 Crossmatch 및 Table 11. Italdata 데이터 혼합 실험 결과로부터 확인하였는데 이는 상이한 센서의 위조지문 데이터 혼합 실험 결과와도 비슷한 맥락으로 데이터베이스 혼합 증강 방법을 적용 시 두 모델의 성능 변화는 반대의 성향을 보인다는 것을 확인하였다.

Table 9. Experimental results of Biometrika data mixing

Sensor / Biometrika	Testing(ACE,%)	
Training	LivDet 2011	LivDet 2013
LivDet 2011	8.85	-
LivDet 2013	-	1.75
LivDet 2011+2013	14.60	1.60

Table 10. Experimental results of Crossmatch data mixing

Sensor / Crossmatch	Testing(ACE,%)	
Training	LivDet 2013	LivDet 2015
LivDet 2013	4.37	-
LivDet 2015	-	2.85
LivDet 2013+2015	5.00	2.21

Table 11. Experimental results of Italdata data mixing

Sensor / Italdata	Testing(ACE,%)	
Training	LivDet 2011	LivDet 2013
LivDet 2011	10.10	-
LivDet 2013	-	1.80
LivDet 2011+2013	14.40	1.25

4.3.3 판별 난이도에 따른 데이터 혼합 실험 결과

앞서 실험한 결과들로부터 위조지문 데이터들을 혼합 시 위조지문 판별에 관한 성능이 변화되는 것을 확인하였다. 또한, 혼합 시 성능 변화에 영향을 주는 요인이 위조지문 데이터의 판별 난이도라 여겨져 판별 난이도에 따른 데이터 혼합 실험을 진행하였다.

실험은 앞서 분류한 Table 2.를 기준으로 판별 난이도가 상대적으로 매우 낮은 데이터베이스 및 높은 데이터베이스를 기준 데이터베이스로 선택한 뒤, 난이도가 낮은 데이터베이스에는 점점 난이도가 높은 데이터베이스를 혼합하는 방향으로 학습 데이터를 구성하였으며 CNN 모델 학습 후 판별률 변화를 확인하였다. 반대로 난이도가 높은 데이터베이스에는 상대적으로 난이도가 낮은 데이터베이스를 혼합하는 방향으로 실험을 진행하였다. 실험 결과 Table 12.와 같이 판별 난이도가 낮은 위조지문 데이터베이스에 점차 난이도가 높은 위조지문 데이터베이스를 혼합하며 CNN 모델을 학습시키는 경우 위조 여부를 판별하는 성능이 향상되는 것을 확인하였으며, Table 13.과 같이 판별 난이도가 높은 위조지문 데이터베이스에 추가적으로 판별 난이도가 낮은 위조지문 데이터베이스를 혼합 할 시 성능이 감소하는 것을 확인하였다.

Table 12. Experimental results of data mixing: adding difficult DB to an easy DB

Level	Testing DB(ACE,%)
Training DB	2
2	6.39
2+4	6.08
2+4+5	4.74
2+4+5+7	4.37
2+4+5+7+9	4.09

Table 13. Experimental results of data mixing: adding easy DB to a difficult DB

Level	Testing DB(ACE,%)
Training DB	9
9	10.10
9+7	11.20
9+7+5	12.05
9+7+5+4	13.25
9+7+5+4+1	16.25

Table 14. Comparison between existing method(affine transformation & horizontal reflection) and data mixing method based on level 1 DB

Training DB Size(level)	Existing method	Data mixing method
2000(1)	1.75	1.75
4000(1,4)	1.80	1.65
6000(1,4,5)	1.75	1.60
8000(1,4,5,7)	1.80	1.60
10000(1,4,5,7,9)	1.75	1.35

Table 15. Comparison between existing method and data mixing method based on level 2 DB

Training DB Size(level)	Existing method	Data mixing method
2000(2)	6.39	6.39
4000(2,4)	6.28	6.08
6000(2,4,5)	4.93	4.74
8000(2,4,5,7)	5.06	4.37
10000(2,4,5,7,9)	4.88	4.09

Table 16. Comparison between existing method and data mixing method based on level 5 DB

Training DB Size(level)	Existing method	Data mixing method
2000(5)	9.70	9.70
4000(5,7)	8.90	8.30
6000(5,7,9)	8.95	7.25

4.3.4 기존 방법과의 성능 비교 실험 결과

기존에 CNN을 활용한 분류에 있어 CNN 모델의 성능을 향상시키기 위해 주로 사용한 방법인 아핀 변환 및 수평 반전을 이용하여 학습데이터를 증강 시켰다. 이렇게 증강된 학습데이터를 활용하여 학습을 시켜 얻은 판별률에 관한 성능과 본 논문에서 제안하는 방법을 적용하여 얻은 성능을 비교한 결과는 Table 14, Table 15, Table 16.과 같다.

성능 비교 실험은 학습에 사용한 학습데이터 크기를 기준으로 삼아 비교하였으며, 비교 결과 본 논문에서 새롭게 제안하는 데이터 혼합 방법이 기존 증강 방법에 비해 더욱 효과적인 방법임을 확인하였다.

V. 결 론

본 논문에서는 CNN 특징을 활용하여 지문 이미지의 위조 여부를 판별하는 알고리즘의 성능을 향상시키기 위한 방법으로 데이터 혼합 증강 방법이라는 기존의 데이터 증강 방법과 다른 개념의 새로운 방법을 제시하였으며, 실험을 통해 제안하는 방법의 타당성을 확인하였다.

실험을 통해 도출한 제안하는 방법 관련 핵심 사항은 위조지문 검출 관련 학습데이터 구성 시 단순하게 데이터의 양을 늘리는 것은 CNN 모델 성능에 향상에 도움이 되지 않으며 오히려 성능 향상에 악영향을 준다는 것이다. 또한 데이터 혼합을 통한 성능 향상을 위해서는 지문 센서의 dpi, 이미지 크기, 위조지문 물질 종류 등과 상관없이 위조 여부 판별을 희망하는 위조지문 데이터베이스보다 상대적으로 판별 난이도가 높은 데이터베이스를 혼합시켜야 한다는 것을 실험을 통해 증명하였다. 이는 판별 난이도가 높은 데이터들을 학습데이터 구성에 추가할 시 생체지문과 위조 지문을 분류하는 CNN 모델을 보다 엄격하게 만들어주기 때문으로 생각된다. 추가적으로 본 논문에서 제안하는 방법은 기존에 주로 사용되었던 증강 방법에 비해 같은 크기의 학습데이터를 사용했음에도 불구하고 보다 향상된 성능을 얻을 수 있음을 실험을 통해 확인하였다.

향후 위조지문 판별률 향상을 위한 학습데이터 구성을 위해 상대적으로 판별 난이도가 높은 위조지문 데이터베이스들만을 모아놓은 하나의 위조지문 데이터베이스 풀(pool)을 구성 할 계획이다. 이러한 풀에 학습을 원하는 데이터베이스만을 추가한 뒤 학습을 시키는 형태로 하여 위조지문 판별에 관한 CNN 모델을 효과적으로 학습시킬 수 있도록 할 시, 학습에 필요한 원본 데이터의 크기가 현저히 줄어들어 생체정보라는 지문의 특성상 데이터를 수집하기 힘들다는 현 문제점을 해결 할 수 있을 것이라 예상된다.

References

〔1〕 J. Wayman, A. Jain, D. Maltoni, "An introduction to biometric authentication systems," Springer London, pp. 1-17, 2005.

〔2〕 Y. N. Shin and M. G. Chun, "Analysis on international financial biometric adoption cases and propose a scheme for Korean financial telebiometrics," Journal of the Korea Institute of Information Security and Cryptology, vol. 25, no. 3, pp. 665-672, 2015.

〔3〕 A. Wiehe, T. Søndrol, Olsen, O. K. and F. Skarderud, "Attacking fingerprint sensors," Gjøvik University College, 2004.

〔4〕 P. Lapsley, J. Lee, D. Pare and N. Hoffman, "Anti-fraud biometric scanner that accurately detects blood flow". US Patent 5,737,439, 1998.

〔5〕 A. Antonelli, R. Cappelli, D. Maio and D. Maltoni, "Fake Finger Detection by Skin Distortion Analysis," Information Forensics and Security, vol. 1, no. 3, pp. 360-373, 2006.

〔6〕 D. Baldisserra, A. Franco, D. Maio and D. Maltoni, "Fake fingerprint detection by odor analysis," in Advances in Biometrics, Berlin Heidelberg, Springer, pp. 265-272, 2005.

〔7〕 A. K. Jain, Y. Chen and M. Demirku, "Pores and ridges: high- resolution fingerprint matching using level 3 features," Pattern Analysis and Machine Intelligence, vol. 29, no. 1, pp. 15-27, 2007.

〔8〕 D. Gragnaniello, G. Poggi, C. Sansone and L. Verdoliva, "Local contrast phase descriptor for fingerprint liveness detection," Pattern Recognition, 9 Jun, 2014.

〔9〕 D. Gragnaniello, G. Poggi, C. Sansone and L. Verdoliva, "Fingerprint Liveness Detection based on Weber Local Image Descriptor," IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications, 2013.

〔10〕 X. Jia, X. Yang, K. Cao, Y. Zang, N. Zhang, R. Dai and J. Tian, "Multi-scale Local Binary Pattern with Filters for

Spoof Fingerprint Detection," Information Sciences, 2013.

[11] L. Ghiani, G. L. Marcialis and F. Roli, "Fingerprint liveness detection by Local Phase Quantization," Proc. IEEE Int. Conf. on Pattern Recognition, 2012.

[12] S. B. Nikam and S. Agarwal, "Local Binary Pattern and wavelet-based spoof fingerprint detection," International Journal if Biometrics, vol. 1, pp. 141-159, 2008.

[13] R.F. Nogueira, R. de Alencar Lotufo, and R.C. Machado, "Evaluating software-based fingerprint liveness detection using convolutional net- works and local binary patterns," IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications, 2014.

[14] C. Wang, K. Li, Z. Wu, Q. Zhao, "A DCNN Based Fingerprint Liveness Detection Algorithm with Voting Strategy," Biometric Recognition, Springer, pp. 241-249, 2015.

[15] R.F. Nogueira, R. de Alencar Lotufo, and R.C. Machado, "Fingerprint Liveness Detection Using Convolutional Neural Networks," IEEE Transactions on Information Forensics and Security, vol. 11, no. 6, pp. 1206-1213, 2016.

[16] Y. Zhang, B. Zhou, H. Wu and C. Wen, "2D fake fingerprint detection based on improved CNN and local descriptors for smart phone," In Chinese Conference on Biometric Recognition, pp. 655-662, 2016.

[17] E. Park, W. Kim, Q. Li, J. Kim, and H. Kim, "Fingerprint liveness detection using patch-based convolutional neural networks," Journal of The Korea Institute of Information Security & Criptology, vol. 27, no. 1, pp. 39-47, 2017.

[18] A. Krizhevsky, I. Sutskever and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," In Advances in neural information processing systems, pp. 1097-1105, 2012.

[19] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A.C. Berg and L. Fei-Fei, "ImageNet Large Scale Visual Recognition Challenge," International Journal of Computer Vision, vol. 115, no. 3, pp. 211-252, 2015.

[20] D. Ciregan, U. Meier and J. Schmidhuber, "Multi-column deep neural networks for image classification," In Computer Vision and Pattern Recognition (CVPR), pp. 3642-3649, 2012.

[21] D. C. Cireşan, U. Meier, J. Masci, L. M. Gambardella and J. Schmidhuber, "High-performance neural networks for visual object classification," arXiv preprint arXiv:1102.0183, 2011.

[22] P. Y. Simard, D. Steinkraus and J. C. Platt, "Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis," ICDAR, vol. no. 3, pp. 958-962, 2003.

[23] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.

[24] W. Kim, Q. Li, E. Park, J. Kim, and H. Kim, "Fingerprint liveness detection and visualization using convolutional neural networks feature," Journal of The Korea Institute of Information Security & Criptology, vol. 26, no. 5, pp. 1259-1267, 2016.

[25] D. Yambay, L. Ghiani, P. Denti, G. L. Marcialis, F. Roli, and S. Schuckers, "LivDet 2011—Fingerprint liveness detection compe- tition 2011, Proc. 5th IAPR Int. Conf. Biometrics (ICB), pp. 208 - 215, 2012.

〔26〕 L. Ghiani et al., "LivDet 2013 fingerprint liveness detection competition 2013," Proc. Int. Conf. Biometrics (ICB), pp. 1-6,2013.
〔27〕 V. Mura, L. Ghiani, G. L. Marcialis, F. Roli, D. A. Yambay and S. A. Schuckers, "LivDet 2015 fingerprint liveness detection competition 2015," Biometrics Theory, Applications and Systems (BTAS), 2015 IEEE 7th International Conference on, Arlington, VA, pp. 1-6, 2015.
〔28〕 D. Gragnaniello, G. Poggi, C. Sansone and L. Verdoliva, "An investigation of local descriptors for biometric spoofing detection," IEEE transactions on information forensics and security, vol. 10, no. 4, pp. 849-863, 2015.

〈저자소개〉

김 원 진 (Weonjin Kim) 학생회원
2015년 2월: 인하대학교 정보통신공학과 졸업
2015년 3월~현재: 인하대학교 정보통신공학과 석사과정
〈관심분야〉 딥러닝, 머신러닝, 바이오인식, 영상처리

김 성 빈 (Cheng-Bin Jin) 학생회원
2009년 6월: 중국 연변대학교 컴퓨터공학과 졸업
2014년 6월: 중국 연변대학교 컴퓨터공학과 석사 졸업
2014년 9월~현재: 인하대학교 정보통신공학과 박사과정
〈관심분야〉 딥러닝, 머신러닝, 행동인식

유 경 송 (Jinsong Liu) 학생회원
2012년 7월: 하남공업대학교 정보공학과 졸업
2015년 6월: 중경우전대학교 정보통신공학과 석사
2016년 9월~현재: 인하대학교 정보통신공학과 박사과정
〈관심분야〉 바이오인식, 영상처리

김 학 일 (Hakil Kim) 종신회원
1983년 2월: 서울대학교 제어계측공학과 졸업
1985년 2월: Purdue Univ. 전기/컴퓨터공학과 석사
1990년 2월: Purdue Univ. 전기/컴퓨터공학과 박사
現 인하대학교 정보통신공학부 교수, (사)바이오인식협의회 회장, IOS/IEC JTC1-Sc37 (Biometrics) 국내 및 국제 전문의원
〈관심분야〉 패턴인식, 컴퓨터비전, 바이오인식, 지능형영상감시