

# 다중 데이터 합성 알고리즘 및 합성곱 신경망을 사용한 신원확인 기법 연구

박지열\*, 옥승호\*\*

## A Study on a Identification Method using Multiple Data Synthesis Algorithms and Convolutional Neural Networks

Ji-Yeol Park\*, Seung-Ho Ok\*\*

---

“본 연구는 과학기술정보통신부 및 정보통신기획평가원의 지역지능화혁신인재양성 (Grand ICT연구센터) 사업의 연구결과로 수행되었음” (IITP-2021-2020-0-01791)

---

### 요 약

본 논문은 다중 생체 데이터 합성 알고리즘 및 합성곱 신경망을 사용한 신원확인 기법을 제안한다. 일반적으로 다중 생체 데이터 기반 신원확인 기법의 경우 다중 네트워크 기반의 멀티모달 딥러닝을 사용한다. 하지만 두 가지 이상의 네트워크를 사용할 경우 연산량이 증가하여 임베디드 시스템에 적용하기 어려운 단점이 있다. 본 논문에서는 얼굴 이미지와 몇 개의 음절로 구성된 인증용 음성을 하나의 데이터로 합성하고 이를 단일 합성곱 신경망에 적용하여 신원확인을 수행하였다. 학습을 위한 실환경 데이터로는 5명의 실험 인원에 대해 각 인원당 얼굴 이미지 8장과 인증용 음성 20개를 조합하여 총 800장의 다중 생체 데이터 합성 이미지를 사용하였다. 실험 결과 제안하는 신원확인 기법은 약 93%의 추론 정확도로 정상 동작함을 확인하였다.

### Abstract

This paper proposes an identification method using multiple biometric data synthesis algorithms and convolutional neural networks. In general, multi-biological data-based identity verification methods use multimodal deep learning. However, computation increases when two or more networks are used, making applying to the embedded system difficult. In this paper, a voice for authentication composed of facial images and several syllables was synthesized into one data, and identification was performed by applying it to one convolutional neural network. As for the actual environmental data for learning, a total of 800 multi-biological synthesized images were used by combining 8 facial images per person and 20 voices for authentication for 5 experimental personnel. As a result of the experiment, it was confirmed that the proposed identification method operates normally with an inference accuracy of about 93%.

### Keywords

identification, face recognition, watermarking, data synthesis, artificial intelligence, CNN

---

\* 동의대학교 인공지능학과 석사과정  
- ORCID: <http://orcid.org/0000-0001-7084-1348>  
\*\* 동의대학교 로봇·자동화공학전공 조교수(교신저자)  
- ORCID: <http://orcid.org/0000-0002-9036-0872>

· Received: Oct. 14, 2021 Revised: Oct. 27, 2021, Accepted: Oct. 30, 2021  
· Corresponding Author: Seung-Ho Ok  
Dept. of Robot·Automation Engineering, Dong-eui University, Korea,  
Tel.: +82-51-890-2265, Email: [osh@deu.ac.kr](mailto:osh@deu.ac.kr)

## 1. 서론

최근 신원확인을 위한 방법으로써 기존의 키패드를 사용한 패스워드 입력 등 접촉식 방식보다 얼굴 및 홍채 인식 등 생체 데이터 기반 비접촉식 인증 기법의 사용이 증가하고 있다. 대표적인 생체 데이터 중 얼굴 이미지를 사용한 딥러닝 기반 얼굴인식 기술의 경우 신원확인 수단으로 널리 활용되고 있으나 사용자의 얼굴 데이터를 위조한 사진만으로도 인증이 되어 보안에 취약한 단점이 있다[1][2].

단일 생체 데이터 인증 방식을 사용하는 신원확인 기법의 문제점을 해결하기 위한 방법으로써 다중 생체 데이터 기반 신원확인 기법이 있다. 일반적으로 두 개 이상의 생체 데이터를 사용하는 다중 생체 데이터 기반 신원확인 기법은 2차 인증 방식을 사용하여 보안을 강화할 수 있다[3]. 멀티모달 딥러닝은 두 개 이상의 딥러닝 네트워크에 다중 생체 데이터를 사용하여 신원확인을 수행한다[4].

이와 관련된 연구로는 얼굴과 의복 데이터를 VGG-FACE 및 VGG-NET으로 구성된 멀티모달 딥러닝에 적용하여 96.27%의 신원 인식 결과를 도출한 연구가 있다[5]. 하지만 멀티모달 딥러닝의 경우 연산량이 많은 두 가지 이상의 네트워크를 사용하기 때문에 소형 임베디드 시스템에 적용하기 어려운 단점이 있다. 이러한 문제를 해결하기 위해 디지털 워터마킹 알고리즘을 사용한 데이터 합성 방법이 있으며, 관련된 연구로써는 홍채와 얼굴 이미지를 합성 후 Radial Basis Function으로 검증하여 96.8% 인식률을 도출한 연구가 있다[6].

본 논문에서 다중 생체 데이터 합성 알고리즘과 합성곱 신경망(CNN, Convolutional Neural Network)을 사용한 신원확인 기법을 제안한다. 제안하는 신원확인 기법은 인증을 위한 생체 데이터로써 얼굴 이미지와 인증용 음성을 사용한다. 인증용 음성은 몇 개의 음절로 이뤄진 단어로써 사용자가 직접 생성한 비밀번호 기능을 수행하며, 데이터 합성 알고리즘을 통해 얼굴 이미지와 합성된다.

본 논문에서는 선행 연구에서 제안한 생체 데이터 합성 알고리즘[7]을 활용하였다. 또한 실환경 데이터 적용을 위해 생체 데이터 합성 알고리즘에서 신원확인 성능 향상을 위해 얼굴과 음성 데이터의

비중을 조정하도록 알고리즘을 개선한다. 본 논문에서는 선행 연구에서 제안한 알고리즘과 개선된 알고리즘에 실환경 데이터를 적용하여 알고리즘별 신원확인 기법을 비교 분석함으로써 기존 연구내용을 확장하였다.

## II. 디지털 워터마킹 알고리즘

### 2.1 디지털 워터마킹 알고리즘 기반 신원확인

디지털 워터마킹 알고리즘은 원작자의 서명과 같은 암호를 이미지 및 영상 등의 저작물에 삽입하여 저작권 보호와 불법복제 방지를 위해 사용된다. 워터마킹 시 원본 데이터의 품질 손상을 최소화하고 데이터 합성을 할 수 있어 다중 생체인식 기술에서 안정적으로 두 가지 이상의 데이터를 합성하여 생체정보를 보호할 수 있다. 워터마킹 알고리즘 기반 다중 생체데이터 합성 기법 연구 중 얼굴 이미지에 지문을 워터마킹한 다중 생체 데이터 합성 기법이 있으며, 실험 결과 워터마크 삽입으로 발생하는 얼굴 이미지 훼손은 얼굴 인식률에 영향을 미치지 않음을 확인하였다[8].

### 2.2 제안하는 신원확인 기법

그림 1은 제안하는 신원확인 기법이다. 데이터 합성 알고리즘에서 얼굴 이미지와 인증용 음성은 주파수 영역으로 변환되며, 얼굴 이미지에 인증용 음성을 합성하여 하나의 합성된 이미지를 만든다. 합성된 이미지는 CNN에 적용하여 학습 및 추론 과정을 수행한다.

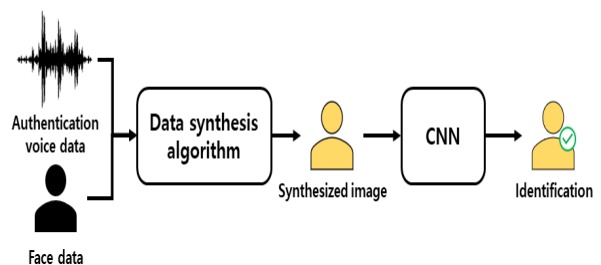
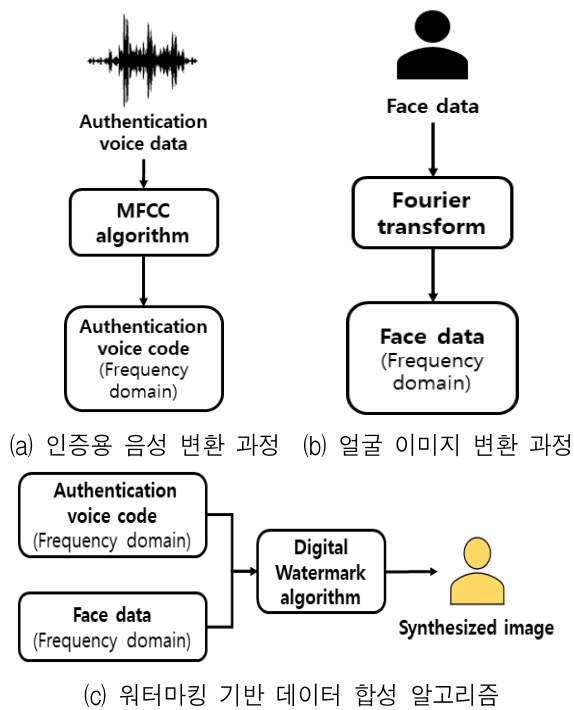


그림 1. 제안하는 신원확인 기법  
Fig. 1. Proposed identification method

### III. 제안하는 다중 생체 데이터 합성 알고리즘

#### 3.1 기존 다중 생체 데이터 합성 알고리즘

그림 2는 선행 연구에서 제안한 선행 연구에서 제안하는 기존 다중 생체 데이터 합성 알고리즘이다. 제안하는 알고리즘은 얼굴 이미지에 인증용 음성을 워터마킹하여 두 개의 생체 데이터를 하나의 이미지로 합성한다. 그림 2(a)는 인증용 음성을 주파수 영역의 인증용 음성 코드로 변환하는 과정이다. 이를 위해 사용되는 MFCC(Mel Frequency Cepstral Coefficient) 알고리즘은 음성 데이터에 mel-filter를 적용해 음성의 고유한 특징을 추출할 수 있다[9]. 그림 2(b)는 얼굴 이미지를 주파수 영역의 얼굴 데이터로 변환하는 과정이며, 푸리에 변환 알고리즘을 통해 주파수 영역 데이터로 변환된다. 그림 2(c)는 그림 2(a)의 인증용 음성 코드와 그림 2(b)의 얼굴 이미지를 디지털 워터마킹 기법을 통해 합성하는 과정이다.



(a) 인증용 음성 변환 과정 (b) 얼굴 이미지 변환 과정  
(c) 워터마킹 기반 데이터 합성 알고리즘  
그림 2. 기존 다중 생체 데이터 합성 알고리즘  
Fig. 2. Existing multiple biometric data synthesis algorithm, (a) Voice for authentication conversion process, (b) Face image conversion process, (c) Watermarking based data synthesis algorithm



(a) 원본 이미지 (b) 합성된 이미지

그림 3. 데이터 합성 결과  
Fig. 3. Data synthesis result, (a) Original image, (b) Synthesis image

변환된 주파수 영역의 인증용 음성 코드는 디지털 워터마킹 알고리즘을 사용하여 주파수 영역의 얼굴 데이터에 합성된다. 합성된 이미지에서 인증용 음성 코드는 특정한 패턴으로 이미지에 출력된다.

그림 3은 기존 다중 생체 알고리즘을 사용하여 얼굴 데이터에 인증용 음성 코드를 합성한 결과이다. 그림 3(a)는 데이터 합성 알고리즘을 수행하기 전 원본 얼굴 이미지이며, 그림 3(b)는 인증용 음성이 합성된 이미지이다. 일반적으로 워터마킹 알고리즘은 저작권 보호를 위해 원본을 훼손하지 않고 워터마크를 삽입한다. 그러나 본 논문에서는 이미지의 특징점에 기반하여 학습 및 추론하는 CNN에 적용하기 위해 인증용 음성의 주파수 값을 증폭하여 인증용 음성 패턴을 가시적으로 나타내었다.

#### 3.2 개선된 다중 생체 데이터 합성 알고리즘

본 논문에서는 얼굴 이미지에 인증용 음성이 합성된 이미지를 CNN에 적용하여 신원확인을 수행한다. 기존 다중 생체 데이터 합성 알고리즘에서 사용한 얼굴 이미지는 그림 3(a) 처럼 흰색 배경이 존재하여 인증용 음성 패턴을 합성된 이미지에 나타낼 수 있다. 그러나 그림 4와 같은 실환경 데이터는 그림 3(a)와 비교하면 흰색 배경의 영역이 적어 그림 5와 같이 합성된 이미지에 인증용 음성 패턴이 그림 3(b)와 다르게 나타난다. CNN은 이미지의 특징점에 기반하여 학습 및 추론을 진행하기 이전에 실환경 데이터를 사용하여 합성된 이미지에서도 사용자 얼굴의 특징점과 인증용 음성의 패턴을 인식할 수 있도록 개선된 다중 생체 데이터 합성 알고리즘을 제안한다.



그림 4. 실환경 데이터 예시  
Fig. 4. Example of actual environmental data



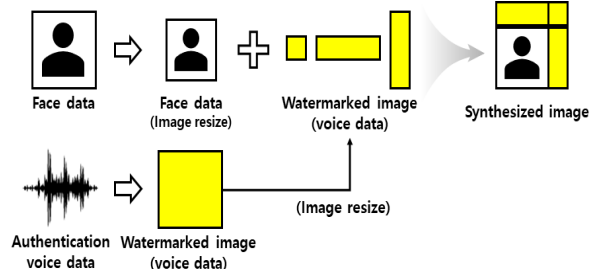
(a) 원본 얼굴 사진 (b) 합성된 이미지  
그림 5. 실환경 데이터 합성 결과

Fig. 5. Actual environmental data synthesis result,  
(a) Original face image, (b) Synthesis image

그림 6은 기존 다중 생체 데이터 합성 알고리즘이며, 원본 얼굴 데이터에 인증용 음성을 합성하였다. 그림 7은 개선된 다중 생체 데이터 합성 알고리즘이다. 그림 7(a)는 동일한 인증용 음성 데이터를 흰색 배경 이미지에 합성 후 3개의 이미지(A, B, C)로 크기를 조정한다. 이를 A, B, C와 같이 얼굴 이미지에 중복되지 않게 배치한 알고리즘으로써 얼굴 이미지 영역을 유지한다. 그림 7(b)는 흰색 배경 이미지에 크기를 줄인 얼굴 이미지를 배치 후 인증용 음성 데이터를 이미지 전체에 합성한 알고리즘으로써 인증용 음성의 비중을 높였다. 그림 6과 7의 데이터 합성 알고리즘은 CNN에 적용하여 성능 평가를 수행하였으며 이를 4.2절에 나타내었다.



그림 6. 기존 데이터 합성 알고리즘  
Fig. 6. Existing data synthesis algorithm



(a) 얼굴 데이터 및 인증 음성 데이터 크기 조절을 통한 데이터 합성 알고리즘



(b) 얼굴 데이터 크기 조절을 통한 데이터 합성 알고리즘  
그림 7. 개선된 다중 생체 데이터 합성 알고리즘

Fig. 7. Improved multiple biometric data synthesis algorithm, (a) Data synthesis algorithm by resized of face data and authentication voice data, (b) Data synthesis algorithm by resized of face data

#### IV. 신원확인 기법 검증 결과

##### 4.1 성능 평가 환경 및 방법

제안하는 신원확인 기법의 성능 평가를 위해 5명의 실험 인원에 대해 실환경 데이터를 사용하였다. 얼굴 이미지는 정면 얼굴 이미지와 마스크, 모자, 선글라스를 착용한 이미지를 사용하였다[10]. 인증용 음성은 실험 인원당 서로 다른 인증용 단어를 녹음한 음성 데이터를 사용하였다.

그림 8과 9는 CNN 학습 및 검증 데이터 조합 과정이다. 그림 8과 9에서  $F_n$ 은 얼굴 이미지를 의미하고  $S_n$ 은 인증용 음성을 의미한다.  $n$ 은  $n$ 번째 실험 인원을 의미한다. 예를 들어  $F_1$  및  $S_1$ 은 각각 첫 번째 인원의 얼굴 이미지와 인증용 음성 데이터를 의미한다. 그림 8은 인원별 학습을 위한 데이터 조합 과정이다. 그림 8의  $F_{n-m}$ 은  $n$ 번째 실험 인원의  $m$ 번째 얼굴을 의미하며 본 논문에서는 실험 인원당 8장의 얼굴 이미지를 사용하였기 때문에  $m$ 의 범위는  $0 < m < 9$ 이다.  $S_{n-m}$ 은  $n$ 번째 실험 인원의  $m$ 번째 인증용 음성을 의미하며 실험 인원당 20개의 인증용 음성을 사용하여  $m$ 의 범위는  $0 < m < 21$ 이다.  $n$ 은 실험 인원의 수를 의미한다.



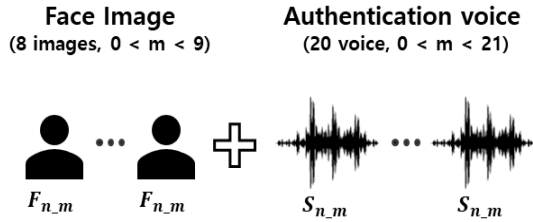


그림 8. 인원별 학습 데이터 조합 과정

Fig. 8. Train data create combination for each person

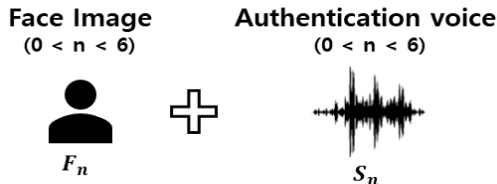


그림 9. 검증 데이터 조합 과정

Fig. 9. Test data combination method

본 논문에서는 실험 인원 5명의 생체 데이터를 사용하기 때문에  $F_{n-m}$ 과  $S_{n-m}$ 에서 n의 범위는 공통으로  $0 < n < 6$ 이다. 예를 들어 그림 10의  $F_{n-8} + S_{n-20}$ 은 n번째 인원의 8번째 얼굴 이미지와 n번째 인원의 20번째 인증용 음성을 합성한 이미지이다. 그림 10은 학습 데이터 조합 결과를 나타내며 학습 데이터는 각 실험 인원당 160장으로써 5명의 실험 인원 에 대한 학습 데이터는 총 800장이다.

그림 11은 그림 9의 검증 데이터 조합 과정을 통해 생성된 검증 데이터의 예시를 나타낸다. 그림 11에서 ①( $F_1 + S_1$ ) 이미지처럼 같은 n 값을 가진 데이터는 정상 인증 여부를 테스트하기 위한 데이터이다. 이와 반대로 그림 11의 ②( $F_1 + S_2$ ), ③( $F_1 + S_3$ ), ④( $F_2 + S_1$ )와 같이 다른 실험 인원의 얼굴 및 음성 데이터가 조합이 된 경우는 오인식률(FAR, False Acceptance Rate)을 테스트하기 위한 데이터이다. 그림 11의 ⑤( $F_1$ )의 경우 인증용 음성이 합성되지 않은 얼굴 데이터로써 인증용 음성이 없는 경우에서의 FAR을 테스트하기 위한 데이터이다.

#### 4.2 검증 결과

제안하는 신원확인 기법을 검증하기 위해 CNN 모델로써 ResNet-18을 사용하였으며 ResNet-18은 경량화된 네트워크로 임베디드 시스템에 적합한 장점이 있다[11].

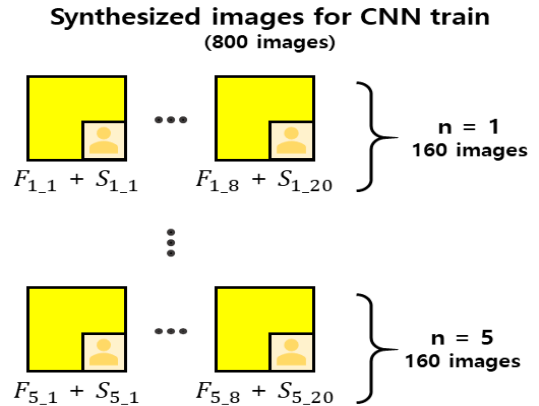


그림 10. 학습 데이터 조합 결과

Fig. 10. Train data combination result

#### Synthesized images for CNN valid (9 images)

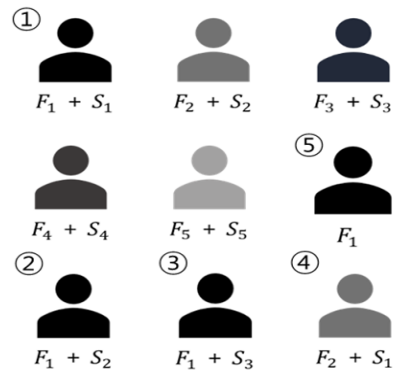


그림 11. 검증 데이터 조합 결과 예시

Fig. 11. Example of test data combination result

딥러닝 프레임 워크로 Tensorflow를 사용하였으며, 학습률(Learning rate)은 0.001, 학습 주기(Epoch)는 100으로 설정하였다. 학습 데이터와 검증 데이터의 비율은 8:2로 나눠 학습하였다.

그림 12는 그림 6과 7의 데이터 합성 알고리즘을 사용하여 합성된 이미지이다. 표 1은 그림 6과 7의 데이터 합성 알고리즘을 그림 11에서 제시한 오인식률 검증을 위한 데이터 ②( $F_1 + S_2$ ), ③( $F_1 + S_3$ ), ④( $F_2 + S_1$ ), ⑤( $F_1$ )를 사용하여 추론된 결과이다. 그림 13은 데이터 합성 알고리즘별 총 4번의 테스트 결과에서 오인식한 검증용 이미지의 비율을 나타낸다. 그림 6의 알고리즘을 사용한 경우 ④( $F_2 + S_1$ )에서 첫 번째 인원으로 추론 결과가 출력되며 오인식률은 25%이다. 이를 통해 그림 6의 기존에 제안된 합성 알고리즘은 인증용 음성의 패턴에 큰 영향을 받는 것을 알 수 있다.



(a) 합성된 이미지 (그림 6) (b) 합성된 이미지 (그림 7.(a)) (c) 합성된 이미지 (그림 7.(b))  
 그림 12. 데이터 합성 기법별 합성 결과  
 Fig. 12. Synthesis result by data synthesis methods,  
 (a) Synthesis image(Fig. 6), (b) Synthesis image(Fig. 7.(a)),  
 (c) Synthesis image(Fig. 7.(b))

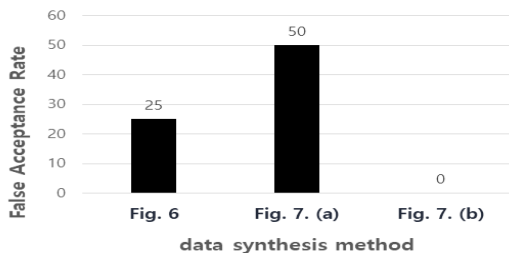


그림 13. 데이터 합성 기법별 오인식률  
 Fig. 13. FAR by data synthesis method

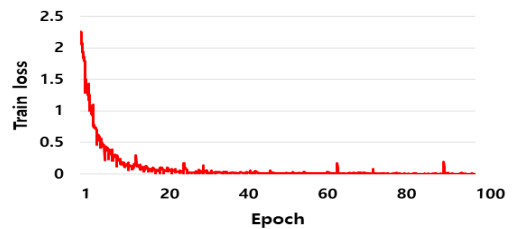
본 논문에서 제안하는 그림 7(a) 알고리즘은 ② ( $F_1 + S_2$ ), ③( $F_1 + S_3$ ) 데이터를 사용한 경우 첫 번째 인원으로 추론되며 오인식률은 50%이다. 이를 통해 (a) 알고리즘은 얼굴 이미지의 특징에 큰 영향을 받는 것을 알 수 있다. 그림 7(b) 알고리즘은 총 4번의 테스트 결과에서 오인식한 검증용 이미지가 없었으며 제안하는 기법이 정상 동작하는 것을 확인하였다. 이에 본 논문에서는 그림 7(b) 알고리즘

표 1. 오인식률 검증을 위한 데이터를 사용한 추론 결과  
 Table 1. Inference results using data for FAR verification

Data synthesis methods	Verification image	Verification result (Experiment number)	FAR
Synthesis algorithm proposed in previous studies (Fig. 6)	$F_1$	none	25 %
	$F_1 + S_2$	none	
	$F_1 + S_3$	none	
	$F_2 + S_1$	The first person(n=1)	
Synthetic algorithm after resize the face and voice size (Fig. 7(a))	$F_1$	none	50 %
	$F_1 + S_2$	The first person(n=1)	
	$F_1 + S_3$	The first person(n=1)	
	$F_2 + S_1$	none	
Synthetic algorithm after resize the face size (Fig. 7(b))	$F_1$	none	0 %
	$F_1 + S_2$	none	
	$F_1 + S_3$	none	
	$F_2 + S_1$	none	

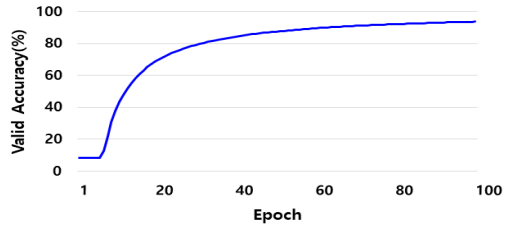
을 사용하여 생체 데이터 합성을 수행하였다.

그림 14는 그림 7(b) 알고리즘을 사용하여 합성된 이미지를 학습한 결과이다. 합성된 이미지는 8:2의 비율을 사용하여 학습 및 검증 데이터셋으로 분류하였다. 그림 14(a)는 학습 데이터셋 기반 학습 손실률을 Epoch에 따라 나타낸 것이며, 그림 14(b)는 검증 데이터셋 기반 추론 정확도를 Epoch에 따라 나타낸 것이다. 실험결과 약 93%의 추론 정확도를 도출하였다.



(a) 학습 데이터셋 기반 학습 손실률

(a) Training loss rate based on training dataset



(b) 검증 데이터셋 기반 추론 정확도

(b) Inference accuracy based on validation dataset

그림 14. 학습 및 검증 데이터셋 기반 실험 결과  
 Fig. 14. Experimental results based on training and validation datasets

동일한 실환경 얼굴 데이터를 사용한 [10]의 신원 인식 네트워크 연구에서는 97%의 추론 정확도를 가진다. 하지만 제안하는 신원확인 기법에서는 ResNet-18을 사용하였으며 [10]의 연구에서는 ResNet-18이 아닌 경량 신원 인식 네트워크를 사용하였다. 서로 다른 네트워크 사용하였기 때문에 추론 정확도의 차이가 발생한다.

표 2는 그림 7(b) 알고리즘을 사용하여 학습한 CNN 모델에 제안하는 신원확인 기법을 테스트한 결과이다. 실험 결과, 표 2에서 정답 값과 일치하는 검증용 이미지는 최대 99.99%의 신원확인 정확도를 보였으며 그 외에 이미지는 최대 27%의 신원확인 정확도를 가진다.

표 2. 검증용 데이터를 사용한 신원확인 테스트 결과  
Table 2. Test result of the identity verification using the data for verification

Answer image (Authentication voice word)	Verification image	Accuracy of inference(%)
$F_1 + S_1$ ("274581")	$F_1 + S_1$	99.99
	$F_1$	0.00
	$F_1 + S_2$	0.00
	$F_1 + S_3$	0.01
	$F_2 + S_1$	0.00
$F_2 + S_2$ ("AI")	$F_2 + S_2$	99.71
	$F_2$	26.38
	$F_2 + S_3$	0.00
	$F_2 + S_4$	0.00
	$F_5 + S_2$	0.00
$F_3 + S_3$ ("Client")	$F_3 + S_3$	99.99
	$F_3$	0.00
	$F_3 + S_2$	0.00
	$F_3 + S_1$	0.00
$F_4 + S_4$ ("0275")	$F_4 + S_4$	99.99
	$F_4$	0.00
	$F_4 + S_1$	27.14
	$F_4 + S_5$	0.00
	$F_3 + S_4$	0.13
$F_5 + S_5$ ("notebook")	$F_5 + S_5$	99.96
	$F_5$	0.00
	$F_5 + S_1$	0.00
	$F_5 + S_2$	0.00
	$F_3 + S_5$	0.00

이를 통해 제안하는 신원확인 기법은 사용자의 얼굴과 인증용 음성이 일치하는 경우에서 신원확인이 정상적으로 동작하는 것을 확인하였으며 서로 다른 생체데이터를 합성한 이미지와 단일 생체데이터에서는 신원확인이 정상적으로 수행되지 않음을 확인하였다.

## V. 결론 및 향후 과제

본 논문에서는 다중 데이터 합성 알고리즘을 사용하여 얼굴 이미지에 인증용 음성을 합성한 이미지를 CNN에 적용한 신원확인 기법을 제안하였다. 제안하는 신원확인 기법은 5명의 얼굴 이미지와 인증용 음성을 조합하여 800장의 합성된 이미지로 학습한다. 학습을 위해 경량화된 네트워크인 ResNet-18을 사용하였으며 학습 결과 93%의 추론 정확도를 가진다.

본 논문에서는 사용자 얼굴 이미지와 인증용 음성이 일치하는 실험에서 최대 99.99%의 신원확인 정확도를 보였으며 서로 다른 생체 데이터를 합성한 이미지에서는 최대 27%의 신원확인 정확도를 가진다. 이를 통해 일치하는 생체 데이터에서만 제안하는 신원확인 기법이 정상 동작함을 확인하였다. 향후 신원확인 테스트를 위한 표준 이미지 및 신원인식률을 높이기 위해 연구된 ArcFace [12] 등의 알고리즘 적용을 통해 제안하는 신원확인 기법의 성능을 개선할 계획이다.

## References

- [1] H. J. Mun and G. H. Kim, "A Survey on Deep Learning based Face Recognition for User Authentication", Journal of Industrial Convergence, Vol. 17, No. 3, pp. 23-29, Sep. 2019. <https://doi.org/10.22678/JIC.2019.17.3.023>.
- [2] W. J. Kim, "Deep learning-based face forgery detection technology trend" Broadcasting and Media Magazine, Vol. 25, No. 2, pp. 52-60, Apr. 2020.
- [3] Y. H. Gil, Y. S. Chung, D.S. Ahn, K. L. Lee,

and S. B. Pan, "Trends of Multi-modal Biometric Technology", Electronics and Telecommunications Trends, Vol. 20, No. 1, pp. 84, Feb. 2005. <https://doi.org/10.22648/ETRI.2005.J.200110>.

[4] L. Yang, D. Jiang, X. Xia, E. Pei, M. C. Ovekenke, and H. Sahli, "Multimodal Measurement of Depression Using Deep Learning Models", 7th Annual Workshop on Audio/Visual Emotion Challenge(AVEC '17), pp. 53-59, Oct. 2017. <https://doi.org/10.1145/3133944.3133948>.

[5] S. E. Hong, W. B. Im, J. W. Park, and H. S. Yang, "Deep CNN-based Person Identification using Facial and Clothing Features", The Institute of Electronics and Information Engineers, pp. 2204-2207, Jun. 2016.

[6] M. Vatsa, R. Singh, P. Mitra, and A. Noore, "Digital watermarking based secure multimodal biometric system", 2004 IEEE International Conference on Systems, Man and Cybernetics, The Hague, Netherlands, Vol 3, Nov. 2004. <https://doi.org/10.1109/ICSMC.2004.1400787>.

[7] J. Y. Park and S. H. Ok, "A Study on CNN based Identification Method using Multiple Data Synthesis", Proceedings of KIIT Conference, Jeju, Korea, pp. 140-143, Jun. 2021.

[8] S. M. Dae, H. J. Seung, H. K. Tae, W. C. Yong, and Y. M. Ki. "An Watermarking Algorithm for Multimodal Biometric Systems", Journal of The Korea Institute of Information Security and Cryptology, Vol. 15, No. 4, pp. 93-100, Aug. 2005.

[9] L. Muda, M. Begam, and I. Ela, "Voice Recognition Algorithms using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques", Journal of Computing, Vol. 2, No. 3, pp. 138-143, Mar. 2010.

[10] H. J. Kwon and S. H. Ok, "Lightweight Identity Recognition Network for Edge Devices", The Institute of Electronics and Information Engineers,

pp. 1769-1771, Jun. 2021.

[11] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition", 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, pp. 770-778, Jun. 2016. <https://doi.org/10.1109/CVPR.2016.90>.

[12] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, "ArcFace: Additive Angular Margin Loss for Deep Face Recognition", 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, pp. 4685-4694, Feb. 2019. <https://doi.org/10.1109/CVPR.2019.00482>.

### 저자소개

박 지 열 (Ji-Yeol Park)



2021년 2월 : 동의대학교  
로봇·자동화공학전공(공학사)  
2021년 3월~현재 : 동의대학교  
인공지능학과(석사과정)  
관심분야 : 딥러닝, 신원확인,  
임베디드 로봇

옥 승 호 (Seung-Ho Ok)



2008년 2월 : 경북대학교  
전자공학과(공학석사)  
2014년 2월 : 경북대학교  
전자공학부(공학박사)  
2014년 3월 ~ 2017년 8월 :  
삼성전자 책임연구원  
2017년 9월 ~ 현재 : 동의대학교  
로봇·자동화공학전공 조교수  
관심분야 : 로봇비전, SoC, VLSI