

## 멀티 채널 이미지를 이용한 손바닥 정보 기반 신원 인식 딥러닝 모델

김범준<sup>0</sup>, 김인기\*,곽정환(교신저자)\*\*

<sup>0</sup>한국교통대학교 컴퓨터공학과,

\*한국교통대학교 교통에너지융합학과,

\*\*한국교통대학교 소프트웨어학과

e-mail: zhskfkseh@gmail.com<sup>0</sup>, cv2@kakao.com\*, jgwak@ut.ac.kr\*\*

### A palm information-based identity recognition deep learning model using a multi-channel image

Beomjun Kim<sup>0</sup>, Inki Kim\*, Jeonghwan Gwak(Corresponding Author)\*\*

<sup>0</sup>Dept. of Computer Engineering, Korea National University of Transportation,

\*Dept. of IT·Energy Convergence, Korea National University of Transportation,

\*\*Dept. of Software, Korea National University of Transportation

#### ● 요약 ●

본 논문에서는 카메라 센서만을 이용하여 손바닥 사진을 촬영하고 추출된 데이터들을 합성하여 멀티 채널 이미지를 생성 및 분류 모델에 입력하여 신원을 확인하는 딥러닝 모델을 제안한다. 이 모델은 손바닥 사진이 입력되면 손바닥 및 손금 세그멘테이션을 이용하여 마스크 이미지를 추출하고 단일 채널로 구성된 이미지들을 멀티 채널 이미지로 합성/재구성하여 신원을 분류하는 딥러닝 모델이다. 이는 카메라 센서 외 다른 센서가 필요 없다는 장점을 가지고 있으며, 비접촉 신원 인식 시스템에 적용할 수 있다.

**키워드:** 손바닥 신원 인식(Hand identification), 멀티 채널 이미지(Multi-channel image)

## I. Introduction

생체 정보를 이용한 신원 인식 기법은 센서 등에 신체의 일부가 접촉하는 접촉식 신원 인식 시스템과 직접적인 접촉없이 신원 인식이 가능한 비접촉식 신원 인식 시스템으로 나뉘며, 최근 COVID-19 사태로 인하여 비접촉식 신원 인식 시스템의 필요성이 두드러지고 있다.

즉, 지문, 정맥 등을 이용할 경우 센서 등에 직접적인 접촉이 된다는 문제점을 가지고 있으며, 이 시스템들은 고가의 센서를 이용하여 데이터를 추출하기 때문에 적지 않은 시스템 설치 비용이 발생한다는 단점을 가지고 있으며, 비접촉식 신원 인식 중 얼굴 인식에 경우 마스크를 벗고 진행해야 한다는 단점을 가지고 있어 현 COVID-19 사태에 적절하지 못한 방법이다.

위 문제들을 해결하기 위해 카메라 센서만을 이용한 손바닥 마스크 이미지 추출 모델, 손금 마스크 이미지 추출 모델, 멀티 채널 이미지 생성 및 멀티 채널 이미지 기반 다중 분류 모델 등을 이용한 손바닥 신원 인식 딥러닝 모델을 제안한다.

## II. Preliminaries

### 1. Related works

#### 1.1 국내 동향

국내에서 손바닥 정보(손금, 지문, 정맥 등)를 활용한 신원 인식 기법은 꾸준히 연구/개발 진행되고 있다. 하지만 대부분의 시스템은 추가적인 센서 등의 의존도가 높게 설계되어 있으며 이는 접촉식 신원 인식의 보안성과 정확도가 비접촉식 신원 인식보다 비교적 높기 때문이다.

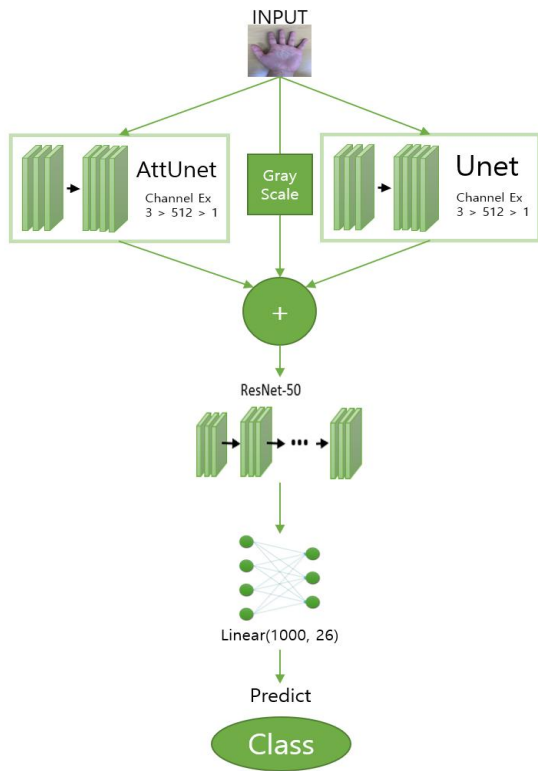


Fig. 1. System Architecture

Fig. 1은 기존의 손바닥 정보를 기반으로한 신원 인식 기법의 문제점을 해결하기 위해 본 논문에서 제안하는 카메라 센서만을 이용한 비접촉식 신원 인식 딥러닝 모델의 아키텍처이며, 연구/개발 환경은 아래 Table 1과 같다.

Table 1. System Environment

Item	Value
CPU	Intel i9-10900K
Memory Size	64GB
GPU	NVIDIA GEFORCE RTX 3090
CUDA Ver	11.1
Python Ver	3.7.11
Pytorch Ver	1.9.1
OS	Ubuntu Server 18.04 LTS

### III. The Proposed Scheme

본 논문에서 제안한 멀티 채널 이미지를 이용한 손바닥 정보 기반 신원 인식 딥러닝 모델은 손바닥 정보 추출을 위한 U-Net[1], Attention U-Net[2]을 사용하였으며, 신원 인식을 위한 다중 분류기 중 일부분이 사전학습된 ResNet50으로 구성되었다.

### IV. Experiment

#### 4.1 Experimental Setting

손바닥 정보를 추출하기 위해 모델 중 U-Net 학습을 위한 데이터

셋은 본인과 지인의 손바닥 사진을 직접 촬영하여 Augmentation을 진행한 결과 총 990장의 데이터 셋을 구성하였으며, Attention U-Net 학습을 위한 데이터 셋은 총 1,039장의 손바닥을 촬영한 “PLSU”[3]를 사용하였다. 각각 Ground Truth로 손바닥과 손금이 Mask된 데이터 셋으로서 데이터 정보는 Table 2와 같다.

Table 2. DataSet INFO

Image size	(224, 224, 3)
Train len	U-Net : 900
	Attention U-Net : 946
Valid len	U-Net : 45
	Attention U-Net : 48
Test len	U-Net : 45
	Attention U-Net : 45

다중 분류 모델을 학습하기 위한 데이터 셋은 본인을 포함한 지인 총 6명의 손바닥 사진을 직접 촬영하여 Augmentation을 진행한 결과 총 780 장의 데이터 셋을 구성하였으며 A부터 F까지 총 6개의 Class를 가지고 있다. Fig. 2는 Training 이미지의 예시이다.

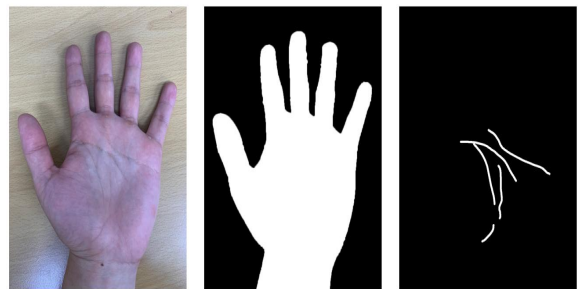


Fig. 2. Example Training Image  
(Left) 원본 이미지 (Middle)손바닥 Mask (Right)손금 Mask

#### 4.2 Multi-channel image

본 논문에서 이야기하는 Multi-channel image란 Gray scale을 적용한 원본 이미지와 모델마다 나온 Mask 예측 이미지를 합성/재구성한 이미지라 정의한다.

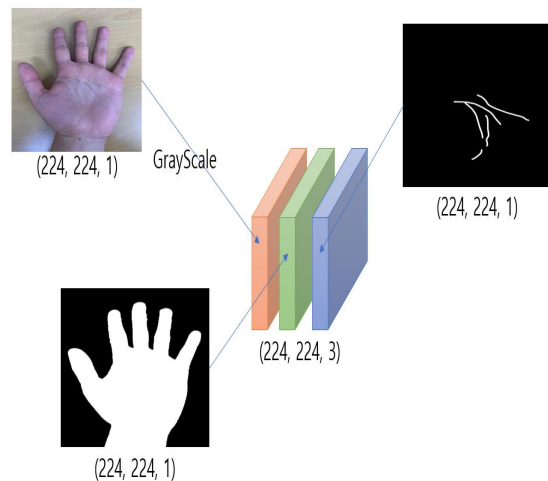


Fig. 3. Multi-Channel Image

상기 Fig 3 처럼 각각의 단일 채널 이미지 배열을 합성하여 RGB 3채널 이미지와 동일하게 구성하여 새로운 이미지 배열을 생성한다.

### 4.3 Results

Figs. 4와 5는 본 논문에서 이야기한 U-Net과 Attention U-Net을 이용하여 손바닥 Mask와 손금 Mask 예측 결과를 보여준다.

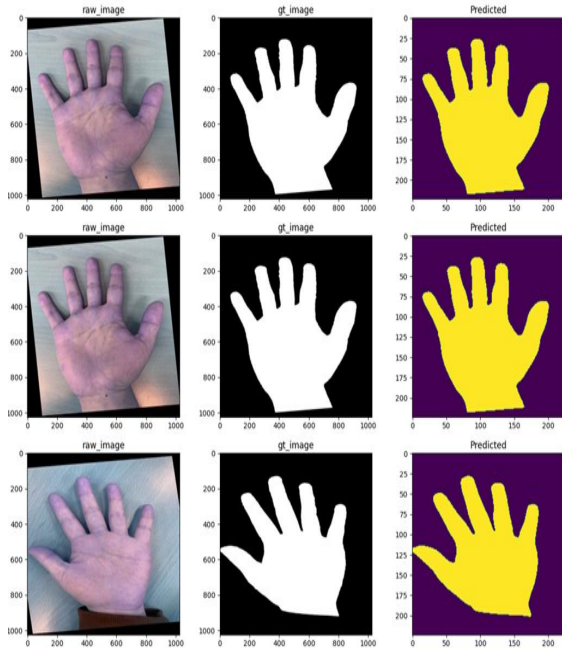


Fig. 4. U-Net Based Hand Segmentation results

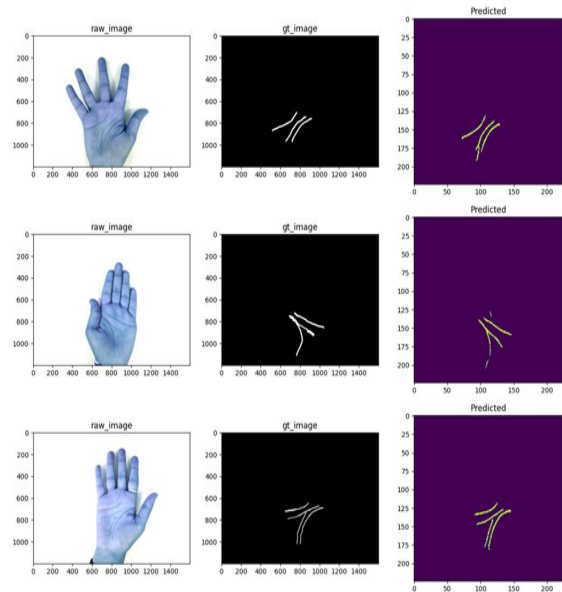


Fig. 5. Attention U-Net Based Palm Lines Segmentation Result

Fig 6은 본 논문에서 제안한 멀티 채널 이미지를 이용한 손바닥 정보 기반 신원 인식 딥러닝 모델의 성능 지표다. 그림처럼 epoch는 10으로 설정하였으며 Loss function은 CrossEntropy를 이용하였다.

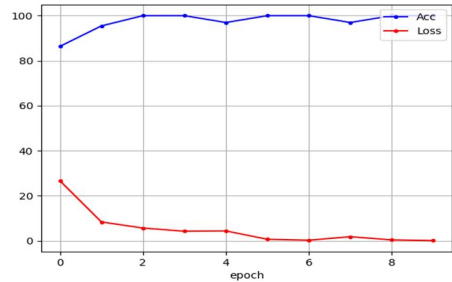


Fig. 6. Model Training Loss And Accuracy Metrics

Figs. 7과 8은 Class 정보와 테스트 데이터셋에 대한 예측 결과를 보여준다.

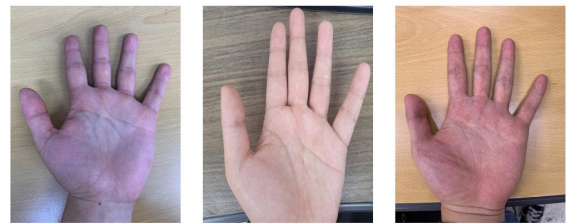


Fig. 7. (Left)class A (Middle)class B (Right)class C

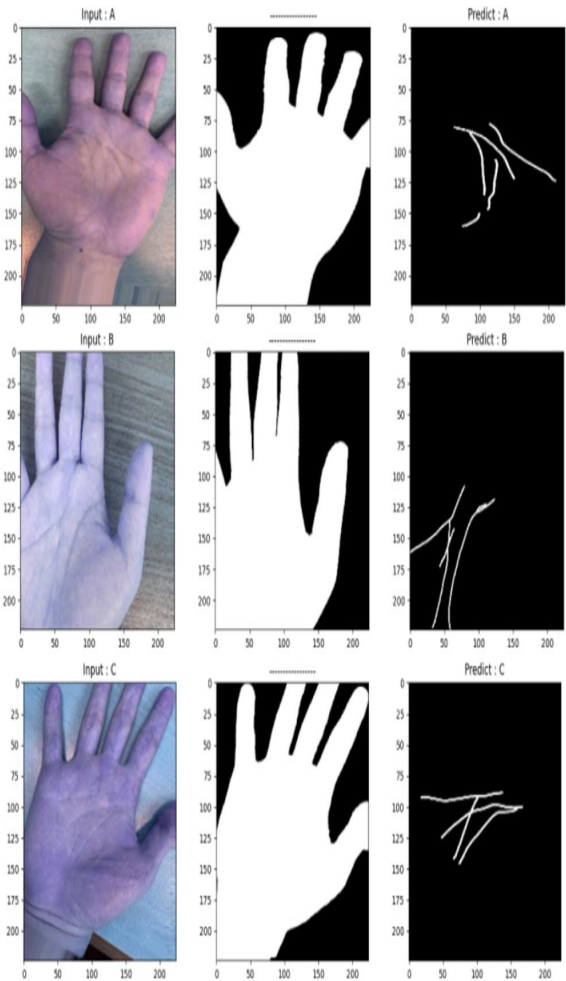


Fig. 8. Model Prediction Results

테스트 데이터 셋에는 원본 이미지에서 손금 정보가 망가지지 않는 한도 내에서 Zoom, Flip, Rotation 등의 이미지 변환을 시킨 데이터 셋이며 예측 점수가 꽤 높은 것을 확인 할 수 있다.

## V. Conclusions

본 연구에서는 카메라 센서만을 이용하여 손바닥 정보를 추출하고 멀티 채널 이미지를 생성하여 생체 인식에 사용 가능한 모델을 제안하였다. 6개의 클래스에 한정해서는 꽤 높은 정확도를 보여주고 있으며, 손바닥이나 손금같은 독립적인 특징들을 멀티 채널 이미지로 재구성하였을 때 효과가 있는 것을 보여준다. 향후 연구 계획은 손바닥 정보를 더욱 세분화하여 3채널이 아닌 그 이상의 채널을 포함한 멀티 채널 이미지 생성에 대한 연구를 진행하며, 이를 바탕으로 한 더욱 많은 사립(클래스)을 분류 가능한 모델을 만들기 위해 현 아키텍처를 계획/연구/수정 단계를 거쳐 비접촉식 생체 인식을 위한 멀티 채널 이미지를 이용한 손바닥 정보 기반 신원 인식 딥러닝 모델을 구현할 계획이다.

## ACKNOWLEDGEMENT

This results was supported by "Regional Innovation Strategy (RIS)" through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(MOE) (2021RIS-001)

## REFERENCES

- [1] O. Ronneberger, P. Fischer, T. Brox "U-Net : Convolutional Networks for Biomedical Image Segmentation", arXiv:1505.04597, 2015.
- [2] O. Oktay, J. Schlemper, L. Le Folgoc, L. Matthew, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Hammerla, B. Kainz, B. Glocker, D. Rueckert "Attention U-Net : Learning Where to Look for the Pancreas", arXiv:1804.03999, 2018.
- [3] Van, T. Pham, Nguyen, S. Trung, Doan, L. Bao, Tran, N. Ngoc, Thanh, T. Minh "Efficient Palm-Line Segmentation with U-Net Context Fusion Module", in 2020 IEEE International Conference on Advanced Computing and Applications(ACOMP), Quy Nhon, Vietnam, 2020, pp. 23-28