**rPPG와 LSTM 방법을 이용한 위조 얼굴 판별 알고리즘**

**Given-Name Surname1, Given-Name Surname 2, Given-Name Surname\*3**

1author's affiliation, Include Department, Institutional address, Country  
2Co-author's affiliation, Include Department, Institutional address, Country

3Co-author's affiliation, Include Department, Institutional address, Country

Corresponding author\* : email@aa.ac.kr

## Abstract

생체인식 기술은 최근 보안시스템 분야에서 널리 사용되고 있으며, 이 중 얼굴인식은 높은 정확도와 실시간 처리가 가능하다는 점에서 각광받고 있다. 하지만 사람의 얼굴은 노출되어 있어 얼굴 영상 등으로 쉽게 위조될 수 있다는 문제가 있어 위조얼굴 검출이 중요한 이슈가 되고있다. 본 논문에서는 rPPG(remote-Photoplethysmogram) 신호와 LSTM을 이용해 얼굴 영상에서 추출한 시계열 영상으로부터 생체신호를 추출하여 실제 얼굴 영상과 얼굴 사진 영상 간의 위조 얼굴 판별 기술을 제안한다. 결과적으로 LSTM 모델에서 필터링 된 생체신호 데이터에 대하여 96.67%의 높은 정확도를 보이는 것을 확인할 수 있다.

**Keywords:** 얼굴인식, 생체신호 검출, 위조 얼굴 검출, rPPG, LSTM, YCbCr

## Introduction

최근 생체정보를 활용한 개인인증기술이 활발히 연구되고 있다. 생체인식 방법은 사람마다의 서로 다른 생체적 특성을 이용하기 때문에 키나 카드와 같은 추가적인 도구를 따로 지니고 다닐 필요가 없으며 공유, 분실, 도난의 염려가 없다. 그리고 인증이 필요할 시 항상 본인이 물리적으로 현장에 있어 사기 행위의 발생 가능성이 적다는 장점으로 인하여 중요한 인증 기술로 인정되고 있다. 하지만 최근 이러한 생체 정보를 위조하여 생체 인식 보안 시스템을 속이는 사례들이 발생하면서 위조 생체 정보 검출에 대한 중요성이 높아지고 있으며 이러한 위조 얼굴 여부를 판단하기 위해 생체 정보의 ‘liveness’를 검출하는 연구들이 진행되고 있다. 본 논문에서는 시계열 영상으로부터 얻어지는 생리신호를 기반으로 사진을 이용한 위조 얼굴을 판별하는 방법을 제안한다. 실제 얼굴에서는 심장박동으로 여겨지는 신호의 주파수 대역이 두드러지게 나타나지만 정지 영상 또는 정지 물체에서는 이러한 대역이 나타나지 않는다는 특성을 이용하는 것이 핵심이다.

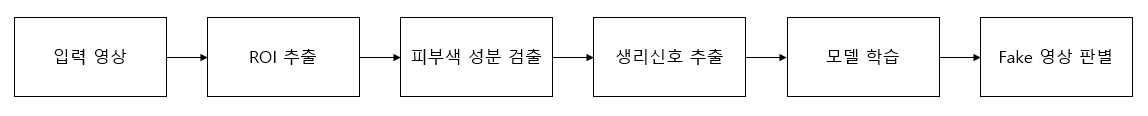
Suh의 “Physiological signal extraction based liveness detection method for face recognition system” 연구에서는 YCbCr 색상 모델에서 영상 기반의 피부색 증폭 방법을 통해 미세 변화를 포착하고, 이로부터 생체신호를 추출하여 현재 얼굴 인식 시스템 앞의 사람이 살아있는지 아닌지를 판별하는 방법을 제안하였다[1]. 실험 결과 생체신호의 변곡점 관점에서 실제 얼굴과 위조 얼굴(얼굴 사진)이 구분되는 결과를 보였다. 그렇지만 충분하지 않은 DB를 가지고 실험하였고, 변곡점 간의 거리 차이만을 보여주었으며 정확성 테스트는 하지 않았다.

Liu의 “3D Mask Face Anti-spoofing with Remote Photoplethysmography” 연구에서는 잡음을 방지하기 위해 Local rPPG신호를 이용하여 생체신호를 추출한다. 여기서 rPPG 신호가 실제 얼굴에서만 일관성을 가지는 특성을 이용하여 주기적 신호인 추정 값 ‘ground truth’와 상호상관 연산을 수행한다. 위 연산을 수행하기 위해서는 ‘ground truth’ 값을 추정하는 과정이 필요하며, 이 추정 값에 의해 성능이 좌우된다.

따라서 본 논문에서는 시계열 영상으로부터 얻어지는 생체신호만을 이용하여 사진을 이용한 위조 얼굴 판별 방법을 제안한다.

## Proposed Work

본 논문에서는 사람의 얼굴 영상으로부터 추출되는 생체신호를 기반으로 LSTM 모델을 학습시켜 실제 얼굴 영상과 위조된 영상을 판별하고자 한다. 다음 그림은 제안하는 방법의 순서도이다.



**Figure 1. 위조 영상 판별 방법 순서도**

먼저 입력 데이터로 시계열 영상을 입력한다. 얼굴의 이동이나 표정 변화는 생체신호와 상관없는 노이즈 성분으로 작용될 수 있으므로 정적인 얼굴 영상을 사용한다. 위조된 영상을 파악하기 위하여 평균 8~10초 길이의 얼굴영상을 사용하여 실험을 진행하였다.

이후 입력된 영상에Viola의 방법을 적용하여 얼굴 영역을 검출한다[3]. 우리는 YCbCr 색상모델을 적용하여 검출된 얼굴 영역을 기반으로 생체신호를 추출하였다. YCbCr 모델에서 Y는 조명 성분, Cb와 Cr은 각각 푸른 색차와 붉은 색차 성분을 나타내는데, 영상에서 Cb와 Cr값의 평균을 통해 피부색의 미세한 변화를 추출하였다. YCbCr 색상 모델을 적용하면 조명 변화를 배제한 색차 성분 관점에서 혈류 흐름에 의한 색 변화를 포착하여 생체신호를 추출할 수 있다[1].

얻은 생체신호에서 영상에 포함된 노이즈 성분을 제거하기위해 bandpass filtering을 적용하여 생체신호와 유사한 주파수 영역대의 신호를 추출하였다. 이후 필터링한 신호에 furrier transform를 적용하여 주파수 영역으로 변환된 신호(Frequency)를 추출하였다.

생체신호를 통해 위조된 얼굴 영상을 검출하기 위해 순환신경망(Recurrent Neural Networks: RNN)의 일종인 Long Short-Term Memory(LSTM)을 이용하였다[4]. LSTM을 이용하면 생체신호의 시계열 관계를 효과적으로 추정하여 위조 얼굴을 효과적으로 검출할 수 있다. LSTM 모델은 4개의 층과 LeakyReLU와 Sigmoid층, 과적합 방지를 위한 배치 정규화와 조기종료 방법으로 구성되어 있으며, 전체적인 모델 구조는 그림과 같다.

테이블이(가) 표시된 사진

자동 생성된 설명

**Figure 2. LSTM 모델 구조**

위 알고리즘을 적용하기 위해서는 피부색의 미세한 변화를 포착할 수 있는 근거리에서 촬영한 최소 10초 이상의 비디오 형태 영상이 필요하다. 따라서 공개된 얼굴 영상 데이터셋은 조건에 부합하지 않아 위조 얼굴 검출 알고리즘에 대한 타당성을 입증하기 불가능하였다. 이에 본 실험에서는 공개 데이터셋이 아닌 별도의 피실험자를 직접 촬영한 영상을 기반으로 실험을 수행하였다. LSTM 모델에 test 데이터를 적용하여 정확도를 도출한 결과 0.97이 나왔으며, F1 score를 도출한 결과 0.97로 위 알고리즘이 위조 얼굴 판별에 있어 좋은 성능을 보인다.

## Conclusion

본 논문은 rPPG와 LSTM 방법을 이용한 가짜 생체신호 판별 알고리즘을 제안한다. 이 작업의 목적은 웹캠 이외에 다른 기기들 없이 영상만을 통해 가짜 생체신호를 판별하는 것이다. 이를 위해 영상에서 생체신호를 추출하고 furrier transform을 통해 Frequency 신호를 획득한다. 가짜 생체신호를 판별하기 위한 알고리즘으로는 LSTM을 이용하며 좋은 성능을 보여줌을 증명하였다.

현재는 ‘실제 얼굴’과 ‘프린트된 얼굴’ 두 가지만을 구분하였지만 이후 연구에서는 replay영상까지 구분할 수 있도록 데이터와 알고리즘을 확장할 계획이다. 또한 생체인식 판별에 걸리는 시간을 단축하여 등록단계 뿐만 아니라 인증, 인식단계에서도 이 시스템을 사용할 수 있도록 하겠다.

## References

[1] Suh, Kun-Ha, and Eui-Chul Lee. "Physiological signal extraction based liveness detection method for face recognition system." Asia-pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology 6.3 (2016): 51-59.

[2] Liu, Siqi, et al. "3D mask face anti-spoofing with remote photoplethysmography." European Conference on Computer Vision. Springer, Cham, 2016.

[3] Viola, Paul, and Michael J. Jones. "Robust real-time face detection." *International journal of computer vision* 57.2 (2004): 137-154

[4] Nishizaki, Hiromitsu, and Koji Makino. "Signal classification using deep learning." *2019 IEEE International Conference on Sensors and Nanotechnology*. IEEE, 2019.