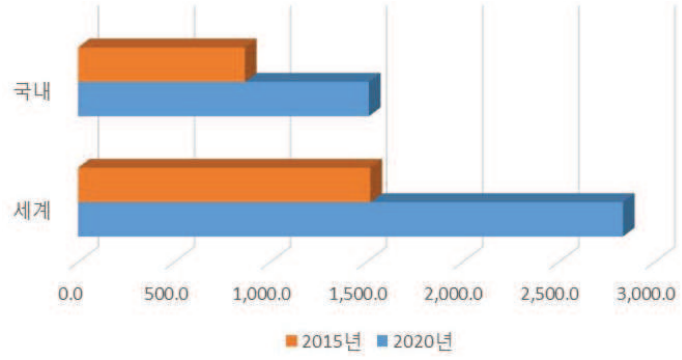
얼굴 인식에서의 스푸핑 공격 탐지 연구 동향

이성훈 조상래* 김수형* 진승헌*
한국전자통신연구원 UST연구생
한국전자통신연구원 책임연구원 *

I. 서론

스마트폰은 우리 삶에 있어서 많은 편리함을 제공해주고 있으며, 이와 더불어 스마트폰에서의 사용자 인증도 중요시되고 있다. 과거 컴퓨터 환경에서 사용되어 왔던 비밀번호 등과 같은 지식기반 인증(knowledge-based authentication)은 작은 스마트폰 기기에서는 사용하기 불편하여 PIN이나 패턴 락(pattern lock)이 사용자의 편의성을 향상시킨 반면, 보안성은 현저히 떨어져서 엿보기 공격 등 다양한 취약점이 존재한다. 스마트폰에 탑재된 카메라의 성능이 향상되면서 얼굴 인식(face recognition) 기술을 이용한 사용자 인증 연구가 활발히 진행되고 있으며, 해당 기술을 탑재한 스마트폰이 출시되고 있다. 얼굴 인식 기반 사용자 인증은 지식 기반 인증의 한계점(주기적인 비밀번호 교체, 비밀번호 입력의 불편함 등)을 극복하여 사용자에게 편의성과 보안성을 제공할 수 있을 것으로



〈자료〉 한국과학기술정보연구원, 2018.

[그림 1] 국내/세계 얼굴 인식 시장 규모

* 본 내용은 이성훈 UST연구생(031-696-3543, backswim@naver.com)에게 문의하시기 바랍니다.
** 본 내용은 필자의 주관적인 의견이며 IITP의 공식적인 입장이 아님을 밝힙니다.
* 이 논문은 2018년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임(No.2016-0-00097, 비대면 본인확인을 위한 바이오 공개키 기반 구조 기술 개발)

[표 1] 애플과 삼성의 얼굴 인식 기반 사용자 인증 기술 비교

구분	애플	삼성
제품명	Face ID	Intelligent Scan
작동 방법	- 30,000개 이상의 보이지 않는 점을 투영 및 분석하여 얼굴 데이터 수집 - 얼굴의 심도맵을 만들고, 적외선 이미지도 촬영 - 심도맵과 적외선 이미지를 수식으로 변환하여 학습 - 등록된 얼굴 데이터와 비교	- 얼굴과 홍채 정보를 함께 사용하여 사용성 및 보안성을 모두 강화한 인증 방식 - 밝은 야외(홍채 인식 어려움)에서는 얼굴 인식으로 보완 - 홍채 인식 시 외모를 자동 업데이트 - 외모 변화가 큰 경우 홍채 매칭 후 얼굴 정보 업데이트
보안 장치	- 인쇄물이나 2D 디지털 사진에는 없는 정보를 통해 얼굴 인식 - 스푸핑 방지(3D 마스크 공격 등) 기능 탑재 - 사용자가 눈을 뜨고 기기를 바라보고 있는지 확인	- 머신 러닝 기반의 홍채 및 얼굴 정보를 동시에 분석하여 스푸핑 공격 방지
성능	- 전체 인구 중 임의의 한 사람이 Face ID를 사용하여 잠금 해제할 확률은 1,000,000분의 1	-

〈자료〉 애플 - Face ID에 적용된 첨단 기술에 관하여, 애플서포트
삼성 - Intelligent Scan(얼굴-홍채 복합 인증)설명, 삼성전자서비스

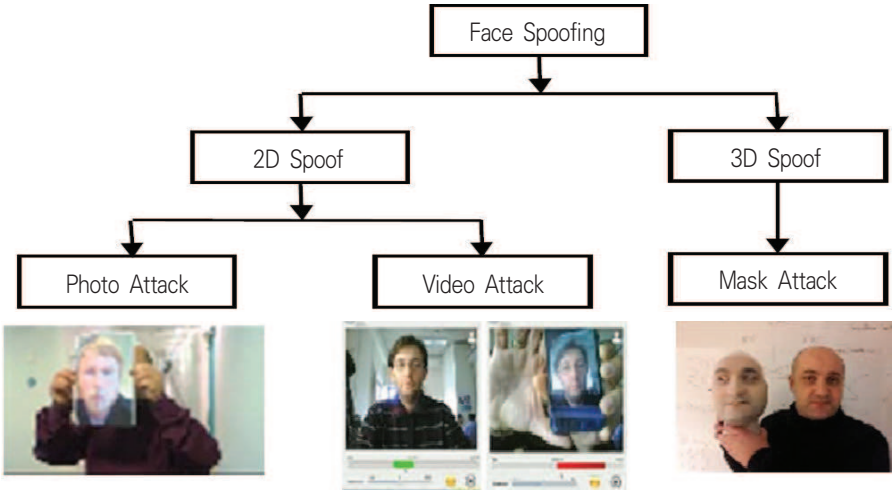
기대되고 있으며, 이에 따라 [그림 1]과 같이 얼굴 인식 시장 규모는 점차 확대되어 2020년까지 국내 연평균 성장률은 11.8%, 세계 연평균 성장률은 13.3%를 기록할 것으로 전망되고 있다[1]. 애플, 삼성과 같은 글로벌 스마트폰 제조업체에서 [표 1]과 같이 자사의 스마트폰에 얼굴 인식 기술을 이용한 사용자 인증 기술을 탑재하여 출시하고 있다[2],[3].

애플의 페이스(Face) ID는 트루뎁스(TrueDepth) 카메라를 이용하여 얼굴 및 적외선 이미지를 촬영한다. 촬영된 이미지는 30,000개 이상의 보이지 않는 점을 특징으로 하여 사용자의 얼굴 패턴을 생성한다. 사용되는 특징점은 스푸핑 방지에도 사용된다. 생성된 이미지는 사용자 얼굴 인증 시마다 비교하여 사용자 본인인지 아닌지를 구분한다[2].

삼성의 얼굴-홍채 복합 인증 솔루션인 인텔리전트 스캔(Intelligent Scan)은 얼굴과 홍채 정보를 모두 사용하여 얼굴 인식이 어려운 경우에 홍채 인식으로 사용자를 인증한다. 또한, 홍채 인식시마다 얼굴 정보를 업데이트하여 가장 최신에 인증된 얼굴 정보를 사용한다는 특징이 있다. 머신 러닝 기반으로 홍채 및 얼굴 정보를 분석하여 스푸핑 공격을 방지한다[3]. 하지만, 출시되자마자 진짜 얼굴이 아닌 얼굴 사진을 이용하거나 가짜 마스크로 얼굴 인식을 속이는 스푸핑 공격으로 인해 보안성에 심각한 문제가 제기되고 있다[4],[5]. 이를 해결하고자 얼굴 인식에서의 안티 스푸핑(anti-spoofing) 연구가 진행되고 있다. 본 고에서는 얼굴 인식에서의 스푸핑 공격을 간략히 설명한 후, 이러한 공격에 대응하기 위한 스푸핑 탐지 기법에 대해 알아보고 향후 연구 방향성에 대해 고찰하고자 한다.

II. 얼굴 인식 스푸핑 공격탐지 기술

얼굴 인식에서의 스푸핑 탐지 기법에 대해 알아보기에 전에 얼굴 인식 스푸핑 공격의 종류에 대해 설명한 후, 이러한 공격들을 탐지하는 기법에 대해 소개한다. 얼굴 인식 스푸핑 공격은 [그림 2]와 같이 인화된 사진 공격(printed photo attack), 비디오 리플레이 공격(video replay attack), 3D 마스크 공격 등 3개의 카테고리로 구분된다.



〈자료〉 A Comparative Study on Face Spoofing Attacks, ICCCA, 2017.

[그림 2] 얼굴 스푸핑 공격 카테고리

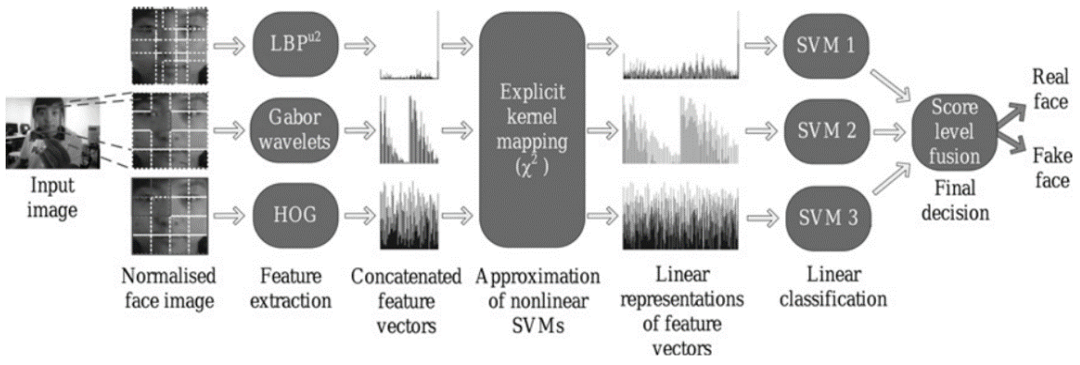
인화된 사진 공격은 얼굴이 등록된 사용자의 사진을 인화하여 카메라에 인식시키는 방법이다. 가장 쉽게 할 수 있는 공격이며, 사진을 약간 왜곡되게(wrap) 구부려서 실제 얼굴처럼 카메라에 인식되게 하여 공격 성공률을 높이는 방법도 있다. 비디오 리플레이 공격은 얼굴이 등록된 사용자의 동영상을 기기(스마트폰, 태블릿, 노트북 등)에서 재생하고 직접 카메라에 보여주어 공격하는 방법이다. 인화된 사진 공격이 단순히 정적인 사진만을 이용한 공격이었다면, 비디오 리플레이 공격은 사용자의 동적인 동영상을 기기에서 재생하여 사용자의 모션(motion)을 고려한 점이 특징이다. 3D 마스크 공격은 한 단계 진화하여 실제 얼굴의 모형을 마스크로 제작하여 공격하는 고급 공격이지만, 마스크 제작에 많은 비용이 들어간다는 단점이 있다. 애플 iPhone X의 얼굴 인식 기반 사용자 인증 솔루션인 페이스(Face) ID에 대해 3D 마스크 공격으로 인증을 성공한 사례가 있는 만큼 강력한 공격이다[4].

이러한 얼굴 인식 스푸핑 공격을 탐지하는 연구는 2015년 이전에는 정적/동적 탐지 위주의 연구가 진행되어 왔으며 인화된 사진 공격에 대한 탐지가 주를 이루었던 반면, 2015년 이후에는 인화된

사진 공격 이외의 다른 공격도 탐지할 수 있도록 공격 탐지 범위가 확장되었고 여러 특징들을 결합하거나 혈류 측정 등 새로운 아이디어를 이용한 탐지 방안 연구가 진행되고 있다. 또한, 딥러닝 기술의 발달로 인해 딥러닝 기술을 탐지 알고리즘에 도입하여 성능 향상을 보여주고 있다.

1. 정적 기반 탐지 기법 - 2015년 이전

2015년 이전의 정적 기반 탐지 기법에 대해 살펴보면, 인화된 사진 공격을 탐지하기 위한 방안으로 인화된 사진과 실제 얼굴과의 차이점을 특징으로 구분한다. 인화된 사진의 경우, 흐릿하거나 희미한 줄이 얼굴과 같이 인화되는 등 실제 얼굴과는 다른 차이점이 있다. 이러한 차이점을 특징으로 하여 실제 얼굴과 인화된 사진을 구분하여 텍스처 분석(texture analysis)을 통해서 확인할 수 있으며 몇 가지 특징들을 활용한다. [그림 3]은 이러한 텍스처 기반 분석을 이용한 스푸핑 탐지 기법의 구성도를 보여준다.



〈자료〉 Face spoofing detection from single images using texture and local shape analysis, IET Biometrics 2012.

[그림 3] 텍스처 기반 얼굴 스푸핑 탐지 구성도

먼저 사용자의 얼굴을 카메라로 입력 받은 후, 해당 이미지에서 특징을 추출할 수 있도록 정규화(normalization)한다. 정규화된 이미지에서 텍스처 분석에서 사용되는 LBP(Local Binary Pattern), Gabor wavelets, HOG(Histogram of Oriented Gradient), DoG(Difference of Gaussians) 등 특징을 추출하고, 기계학습 알고리즘의 한 종류인 SVM(Support Vector Machine)을 이용하여 특징별로 정상 사용자의 얼굴 패턴을 학습한다. 학습된 모델은 추후 정상 사용자의 실제 얼굴인지 아니면, 인화된 사진을 이용한 스푸핑 공격인지를 탐지하는데 사용된다. LBP는 컴퓨터 비전 분류에서 사용되는 시각적 구분자의 한 종류로서 질감을 분류하는데 있어 효과적이며 많이 사용되는 특징 중 하나이고, 물체 식별력이 높고 조명 변화에 강인하고, 연산이 간단하다는 장점이 있다[6]. Gabor

wavelets는 얼굴 이미지의 질감 표현을 향상시키는 특징으로 기본 아이디어는 다양한 스케일과 오리엔테이션 관점에서 특징을 추출하는 것이다. 해당 특징은 이미지의 다양한 조명에 대해 안정성을 제공하며 이미지의 왜곡, 회전, 크기, 변형에 안정적이다[7]. HOG는 특정 부분에 대한 밝기의 분포 방향을 히스토그램으로 표현하여 특징 벡터로 사용한다. 특히, 물체의 모양을 표현하는데 적합한 것으로 알려져 있다[8]. DoG는 이미지에서 고주파 성분을 추출하여 특징으로 사용하며 실제 얼굴과 인화된 사진을 구분한다[9].

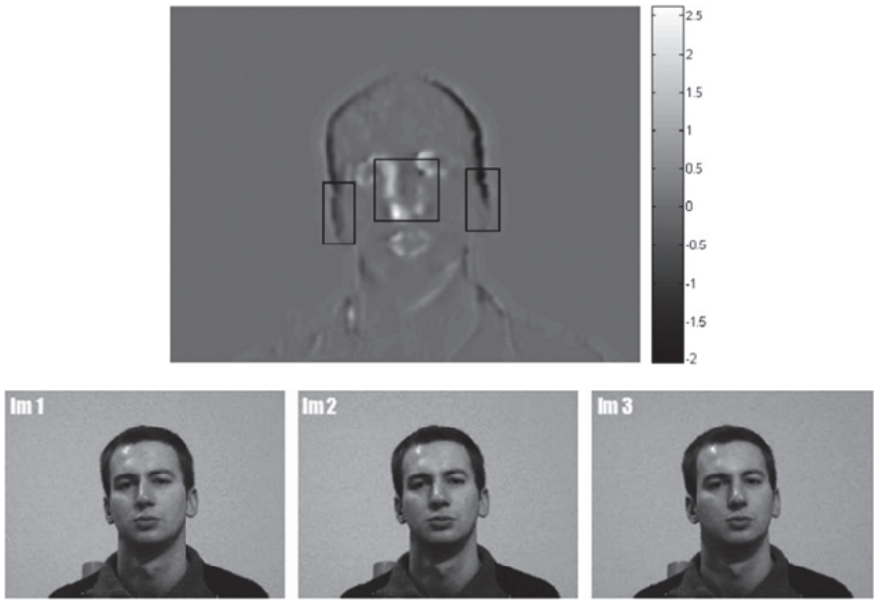
카메라로 인식되어 정규화된 이미지로부터 이러한 특징들을 추출한 이후, 각각 SVM 알고리즘을 이용한 분류를 통해 분류 결과를 점수로 산출하고 산출된 점수들을 결합하여 최종 분류하고, 실제 사용자의 얼굴 또는 인화된 사진으로 구분한다. SVM은 기계학습에서 많이 사용되는 알고리즘 중 하나로 두 개의 카테고리 중 어느 하나에 속한 데이터의 집합이 주어졌을 때, 주어진 데이터 집합을 바탕으로 하여 새로운 데이터가 어느 카테고리에 속할지 판단하는 비확률적 이진 선형 분류 모델이다[10].

2. 동적 기반 탐지 기법 – 2015년 이전

동적 탐지 기법의 경우, 정적 탐지와 마찬가지로 실제 얼굴과 인화된 사진 공격을 탐지하는 것을 목표로 한다. 사진과 달리 실제 얼굴은 미묘한 움직임 및 색상 변화를 통해 실제 사용자의 얼굴인지 확인할 수 있다. 예를 들어, 얼굴의 특정 부위(중앙 부위의 코, 바깥 부위의 귀 등)의 특징점 및 라인 변화, 눈 깜빡임, 배경 색의 차이 등을 이용한다. 하지만 탐지 범위가 인화된 사진만 탐지가 가능하고, 비디오 리플레이 공격과 마스크 공격은 탐지할 수 없다는 한계가 있다.

얼굴의 특정 부위의 특징점 및 라인 변화를 통한 탐지 기법은 OFL(Optical Flow of Lines)을 통해 얼굴 움직임을 검출한다. [그림 4]는 실제 사용자가 얼굴을 인식할 때, 얼굴 움직임의 변화에 따른 OFL 탐지 부분을 보여주며, 얼굴의 중앙 부분과 바깥 부분의 변화를 감지한다. OFL은 특징점과 라인의 움직임 차이를 이용하여 탐지한다. 실제 사용자의 얼굴은 좌우로 움직일 때마다 코(중앙 부분), 귀(바깥 부분)의 라인과 점이 바뀌지만, 인화된 사진은 라인과 점이 항상 고정되어 있어 탐지가 가능하다[11].

또 다른 기법으로 눈 깜빡임(eye blink)을 이용하여 실제 사용자의 얼굴과 인화된 사진을 구분하는 탐지 연구도 있다. 보통 사람의 경우, 눈 깜빡임은 1분당 15~30번 정도이고, 한 번의 눈 깜빡임은 250 밀리세크(milliseconds) 정도이다. 15 프레임(frame)으로 촬영할 경우, 프레임간 인터벌(interval)은 70 밀리세크 이하이며, 즉 두 개 이상의 프레임에서 눈 깜빡임을 확인할 수 있다.



〈자료〉 Non-intrusive liveness detection by face images, Image and Vision Computing, 2009.

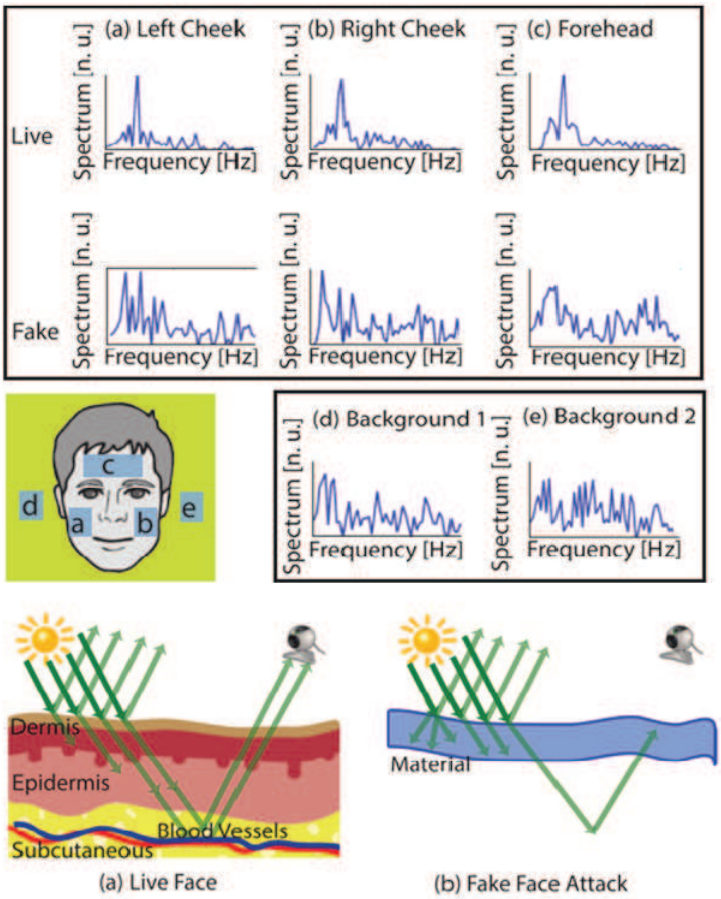
[그림 4] 실제 얼굴의 미묘한 움직임을 탐지하여 인화된 사진 공격 탐지

장점으로는 사용자의 인지 및 추가 행동이 필요 없고, 추가 하드웨어가 필요 없다[12]. 이 외에도 얼굴과 배경 간의 motion coefficient 분석을 통해 실제 사용자와 인화된 사진을 구분하는 연구도 있다[13].

앞서 살펴본 동적/정적 탐지 연구는 인화된 사진 공격 탐지에 국한되어 있다는 한계가 있다. 또한, 실험을 위해 사용한 DB(Database)도 최근 스푸핑 공격이 활발히 이루어지고 있는 스마트폰에서 수집한 데이터가 아닌 일반 웹 카메라나 노트북의 웹캠을 통해 수집한 데이터이며, 이러한 데이터로 실험한 결과이기 때문에 스마트폰을 이용한 스푸핑 공격을 탐지함에 있어 좋은 성능을 기대하기 어려울 수 있다.

3. 복합적 탐지 기법 - 2015년 이후

2015년 이후 딥러닝을 이용하여 탐지율을 향상시키며 다양한 스푸핑 공격을 탐지할 수 있도록 목표 범위를 확장하였다[14]. 또한, 동적과 정적 탐지에서 사용하는 특징들을 복합적으로 결합하거나 혈류 측정과 같은 아이디어를 이용하여 다양한 측면에서 탐지하는 연구가 진행되고 있다. 공격 탐지 범위는 2015년 이전의 연구가 인화된 사진을 이용한 공격에만 국한되어 있었다면, 2015년 이후에는 인화된 사진 공격, 비디오 리플레이 공격, 3D 마스크 공격 등을 모두 탐지할 수 있도록



〈자료〉 PPGSecure: Biometric Presnetation Attack Detection Using Photopletysmograms, IEEE FG 2017, 2017.

[그림 5] PPG를 이용한 실제 얼굴과 스푸핑 공격의 차이점 확인

탐지 범위를 확장하고 있다는 점도 중요한 특징이다. 또한, 탐지를 위한 분류 알고리즘도 딥러닝 기법을 도입하여 기존의 분류 알고리즘보다 향상된 탐지 성능을 보이고 있다.

기존의 동적, 정적에서 사용하던 특징 이외에도 PPG(Photoplethysmogram)를 이용하여 실제 사용자의 얼굴과 스푸핑 공격을 탐지하는 연구도 있다[15],[16]. PPG는 실제 사람의 얼굴과 비디오 녹화 영상으로부터 혈류를 확인하고, 혈류로 인한 색의 변화를 감지하여 실제 얼굴과 스푸핑 공격을 구분한다. 실제 사람의 얼굴과 달리 동영상이나 인화된 사진의 경우 혈관이 보이지 않기 때문에, [그림 5]와 같이 색의 변화가 실제 사람의 얼굴과 다르게 나타난다.

정적, 동적 특징들을 복합적으로 결합하여 탐지하는 경우, 인화된 사진 공격, 비디오 리플레이 공격, 3D 마스크 공격 등을 탐지할 수 있다[17]. 기존 연구에서 사용하던 특징들을 어떻게 결합하느냐에 따라 공격 탐지 범위가 확장됨을 보였고, 분류 알고리즘보다는 특징 결합이 중요함을 시사하고

있다. 분류 알고리즘으로 SVM과 딥러닝 기법을 이용한 뉴럴 네트워크(기계학습과 인지과학에서 인간의 뇌와 같이 생물학의 신경망에서 영감을 얻은 통계학적 학습 알고리즘)[16]를 비교 실험하여 뉴럴 네트워크 사용 시에 SVM보다 약간의 성능 향상을 보였다. 하지만, 해당 실험은 컴퓨팅 파워가 충분한 데스크톱 환경에서 진행되어, 최근 스마트폰에서 얼굴 인식 기반 사용자 인증 기술이 많이 사용되고 있는 만큼 스마트폰 환경에서의 실험 결과가 없다는 점이 아쉬운 부분이다. 딥러닝은 연산 처리 과정이 복잡하여 강력한 컴퓨팅 파워가 요구되는 만큼 스마트폰에 탑재하여 사용하기 어려우며, 스마트폰에서 작동될 경우 경량화가 필요하다.

III. 결론 및 시사점

스마트폰에서 얼굴 인식 기반 사용자 인증은 2D 얼굴 인식에서 카메라의 성능 향상 및 적외선 카메라와 3D 센서 탑재 등을 통해서 3D로 얼굴을 인식하여 사용자 인증의 성능을 강화하는 추세이다. 이에 따라 스푸핑 공격은 기존의 인화된 사진 공격 및 비디오 리플레이 공격보다는 3D 마스크 공격에 의한 스푸핑 공격이 많이 발생할 것으로 보인다. 2017년에 출시된 애플의 iPhone X의 경우 적외선 카메라를 이용하여 사용자의 얼굴 형태를 인식하는 기술을 도입했음에도 불구하고 3D 마스크 공격으로 인해 스푸핑에 성공한 사례가 있는 만큼, 3D 마스크 공격은 단순히 카메라 하나만을 이용하여 탐지하기에는 한계가 있다. 즉, 카메라 이외에 실제 얼굴을 인식할 수 있는 추가 하드웨어가 필요할 것으로 보이며 이렇게 추가된 하드웨어를 이용하여 어떻게 3D 마스크 공격을 탐지할 것인지에 대한 연구가 중요할 것으로 보인다.

삼성, 애플, 화웨이 등 글로벌 스마트폰 제조업체들은 3D 센서와 같은 다양한 센서 탑재 및 카메라의 성능 향상을 통해 3D 얼굴 인식 기반 사용자 인증 기술을 개발하여 차세대 스마트폰에 이미 도입했거나 도입할 예정이다[19]. 시장조사업체 카운터포인트리서치에 의하면, 2020년 전세계 스마트폰의 64%가 얼굴 인식 기반 사용자 인증 기술을 탑재할 전망이다[20]. 이는 2017년 기준

[표 2] 현재와 향후의 얼굴 인식 스푸핑 탐지 비교

구분	현재	미래
탐지 하드웨어	카메라로 촬영된 영상만 활용	적외선 카메라, 3D 센서 등 추가 하드웨어 활용
분류 알고리즘	SVM 등 기계학습 알고리즘	딥러닝 계열 기계학습 알고리즘
탐지 범위	인화된 사진, 비디오 리플레이 공격	3D 마스크 공격 및 현재의 공격 포함

〈자료〉 한국전자통신연구원 자체 작성

스마트폰의 얼굴 인식 탑재 비중이 5%였던 것과 비교하면 대폭 늘어난 수준으로 스푸핑 탐지 기술 확보가 더욱 강조될 것이다.

또한, 분류 알고리즘도 딥러닝이 여러 분야에서 좋은 성능을 보이고 있는 만큼 스푸핑 탐지 분야에서도 딥러닝의 활용이 좋은 탐지 결과를 보일 것으로 기대된다. 다만, 딥러닝은 상당한 연산처리를 요구함에 따라 현재의 스마트폰 컴퓨팅 파워로는 딥러닝을 이용한 분류 알고리즘의 연산처리는 어렵다. 이에 기업 및 학계에서 모바일 기기용 AI 반도체 연구에 박차를 가하고 있으며[21], 향후 스마트폰에서도 딥러닝을 비롯한 기계학습 모델을 실행할 수 있을 것으로 보인다. 이에 따라 앞으로의 스푸핑 공격 탐지 연구에서는 스마트폰에 탑재되는 다양한 센서와 향상된 카메라 기능과 함께 딥러닝을 활용하여 좋은 성능을 기대할 수 있을 것으로 기대된다.

[참고문헌]

[1] MarketsAndMarkets, "Next Generation Biometrics Market-Global Forecast to 2020," 2015.

[2] https://support.apple.com/ko-kr/HT208108

[3] https://www.samsungsvc.co.kr/online/faqView.do?faqId=KNOW0000039287

[4] A. Greenberg, "Hackers say they've broken face id a week after iphone X release," WIRED, 2017. 11. 12.

[5] 유진상, "갤럭시 S8 얼굴 사진으로도 잠금 풀려 유튜브 동영상 논란", IT조선, 2017. 3. 31.

[6] 임길택, 원철호, "아다부스트 학습과 비정방형 Differential LBP를 이용한 얼굴영상 특징분석", 한국멀티미디어학회논문지, 19.6, 1014-1023, 2016. 6.

[7] 전인자, 정경용, 이영호, "의료자산보험에서 얼굴인식을 위한 가보 웨이블릿 분석", 한국콘텐츠학회논문지, 11.11, 10-18, 2011. 11.

[8] 최미순, 이정환, 노태문, 심재창, "HOG 특징 및 영상분할을 이용한 부스팅분류 기반 자동차 검출 기법", 정보과학회논문지, 16.10, 955-961, 2010. 10.

[9] Z. Zhiwei et al., "A Face Antispoofing Database with Diverse Attacks," Proceedings of IAPR Int. Conf. Biometrics(ICB), 2012.

[10] https://en.wikipedia.org/wiki/Support_vector_machine

[11] K. Kollreider, H. Fronthaler, and J. Bigun, "Non-intrusive Liveness Detection by Face Images," Image and Vision Computing, 27.3, 233-244, 2009.

[12] P. Gang et al., "Eyeblink-based Anti-spoofing in Face Recognition from a Generic Webcamera," IEEE 11th International Conference on Computer Vision, 2007.

[13] Andre Anjos, and Sebastien Marcel, "Counter-measures to photo attacks in face recognition a public database and a baseline," International Joint Conference on Biometrics(IJCB), 2011.

[14] Di Wen, Hu Han, and Anil K. Jain, "Face Spoof Detection with Image Distortion Analysis," IEEE Transactions on Information Forensics and Security, 10.4, 746-761, 2015.

[15] Ewa Magdalena Nowara, Ashutosh Sabharwal, and Ashok Veeraraghavan, "PPGSecure: Biometric

Presentation Attack Detection Using Photopleyhsmograms," IEEE International Conference on Automatic Face&Gesture Recognition, 2017.
[16] L. Siqi et al., "3D Mask Face Anti-spoofing with Remote Photoplethysmography," European Conference on Computer Vision(ECCV), 2016.
[17] F. Litong et al., "Integration of Image Quality and Notion Cues for Face Anti-spoofing: A Neural Network Approach," Journal of Visual Communication and Image Representation, 38, 451-460, 2016.
[18] https://en.wikipedia.org/wiki/Artificial_neural_network
[19] 이은정, "스마트폰 생체인식 진화 3D 안면인식 활짝," 지디넷코리아, 2018.02.21
[20] Counterpoint, "More than one billion smartphones to feature facial recognition in 2020," 2018. 2. 7.
[21] 김지은, 이종희, "마윈도 뛰어든 AI반도체 국내기술 수준은," 뉴시스, 2018. 4. 23.