머클트리를 활용한 영상무결성 검사 기법

[1*]강윤희 [2]장은영 [3]권태언

Video Integrity Checking Scheme by Using Merkle Tree

[1*] Yun-Hee Kang, [2] Eun-Young CHANG, [3] Taeun Kwonk

요 약

최근 다양한 분야에서 영상과 사운드를 포함한 디지털 콘텐츠가 생성되어 인터넷을 통해 클라우드에 전송된 후 저장되어 활용되고 있다. 디지털 콘텐츠의 활용을 위해서는 해당 데이터 무결성(data integrity) 검증은 필수적이며, 검증 자료의 네트워크 대역폭 효율성 보장이 필요하다. 이 논문에서는 영상데이터의 무결성 검증을 위한 데이터들을 유지 및 관리하며 제공하는 서버의 설계 및 구현에 관하여 기술한다. 서버는 영상데이터를 획득하는 모듈인 Logger로부터 영상데이터를 전달받아 저장하며, 영상데이터의 검증을 수행하는 모듈인 Verifier에 검증에 필요한 데이터를 제공하는 기능을 수행한다. 이후 해시값을 사용하여 경량 머클트리를 구성한다. 경량 머클트리(light-weight Merkle tree)는 두 버젼의 영상프레임 인덱스의 해당 영상프레임 변경사항을 개별 해시값의 비교 없이도 빠르게 무결성 위반을 검출할 수 있다. 이를 위해 네트워크 대역폭 효율성을 갖도록 디지털 콘텐츠의 해시값을 생성하여 경량 머클트리를 구성하고, 이를 무결성 검증의 증명 수행 결과로 제시한다.

Abstract

Recently, digital contents including video and sound are created in various fields, transmitted to the cloud through the Internet, and then stored and used. In order to utilize digital content, it is essential to verify data integrity, and it is necessary to ensure network bandwidth efficiency of verified data. This paper describes the design and implementation of a server that maintains, manages, and provides data for verifying the integrity of video data. The server receives and stores image data from Logger, a module that acquires image data, and performs a function of providing data necessary for verification to Verifier, a module that verifies image data. Then, a lightweight Merkle tree is constructed using the hash value. The light-weight Merkle tree can quickly detect integrity violations without comparing individual hash values of the corresponding video frame changes of the video frame indexes of the two versions. A lightweight Merkle tree is constructed by generating a hash value of digital content so as to have network bandwidth efficiency, and the result of performing proof of integrity verification is presented.

Keywords: *Digital contents, Data integrity, Authenticity verification, Hash value, Light-weight Merkle tree*

[1*] Corresponding Author 백석대학교 교수 (yhkang@bu.ac.kr)
[2] 국립공주대학교 교수 (ceyng@kongju.ac.kr)
[3] ㈜하스퍼 수석연구원 (peterkwon@harsper.co.kr)

Received: Dec. 04, 2022, Revised: Dec. 16, 2022, Accepted: Dec. 20, 2022

I. 서론

스마트 폰, CCTV, 센서가 연결된 사물인터넷 환경의 출현으로 인터넷에 연결된 장치로부터 생성되는 디지털 콘텐츠들이 빠르게 증가하고 있으며 이를 활용한 응용 요구가 커지고 있다. 데이터의 활용 분야는 지속적으로 넓어지고, 그에 따라 데이터 시장 또한 확장하고 있다[1][2]. 이와 더불어 데이터에 대한 변질 혹은 제삼자의 개입으로 수정된 데이터 또한 많아지고 있다. 최근 다양한 분야에서 영상과 사운드를 포함한 디지털 콘텐츠가 생성되어 인터넷을 통해 클라우드에 전송된 후 저장되어 활용되고 있다. 이들 디지털 영상 콘텐츠는 디지털 자료의 특성상 복제 및 수정이 용이하며, 내용 변경으로 인해 데이터 무결성 위반을 발생할 수 있다[3][4]. 올바른 디지털 콘텐츠의 활용을 위해 해당 데이터 무결성(data integrity) 검증은 필수적이며, 검증 자료의 네트워크 대역폭 효율성 보장이 필요하다. 그러나 다수의 상업용 영상 획득 및 저장시스템은 위변조와 삭제 등의 내용 수정을 확인하는 무결성 검증 기능을 제공하지 못한다. Figure 1(a) 와 Figure 1(b)는 정상 영상과 훼손된 영상의 예를 보인다.

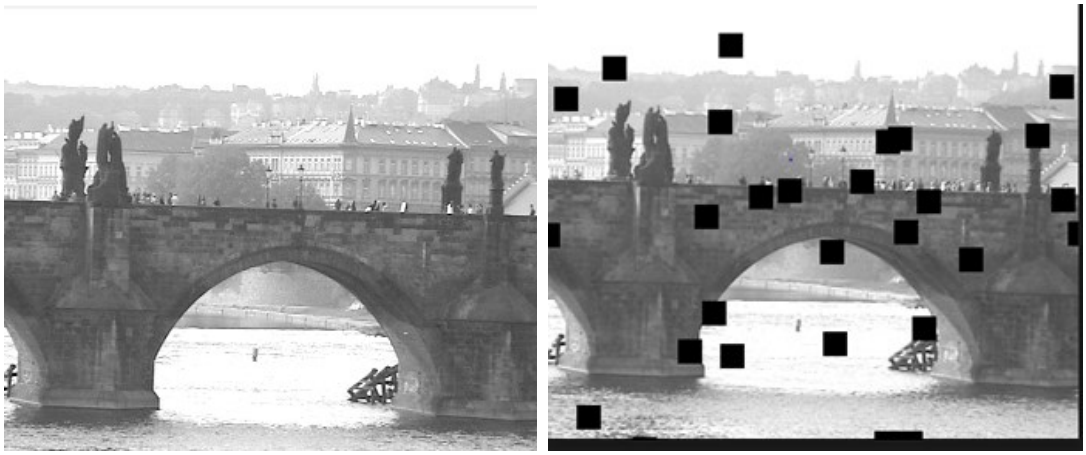

Figure 1. (a). Normal video frame, (b). Tainted video frame
그림 1. (a) 정상 영상 프레임, (b) 훼손된 영상 프레임

데이터 무결성은 해당 디지털 콘텐츠의 정확성과 일관성을 유지하고 보증하는 것으로 데이터 무결성의 유지를 위해서는 허가 받지 않은 사용자에 의해 데이터가 수정되었는지를 감지하는 기능이 필요하다. 데이터 무결성을 보장하기 위해 블록체인을 기반으로 한 영상 무결성 기법이 제안되었으며, 데이터를 통하여 의사결정을 하는 컴퓨터 포렌식과 같은 분야에서는 데이터의 무결성과 데이터 생산 과정에서의 진본 확인은 필수적이다[3][4][5].

이 논문에서는 영상데이터의 무결성 검증을 위한 데이터들을 유지 및 관리하며 제공하는 서버의 설계 및 구현에 관하여 기술한다. 서버는 영상데이터를 획득하는 모듈인 Logger 로부터 영상데이터를 전달받아 저장하며, 영상데이터의 검증을 수행하는 모듈인 Verifier 에 검증에 필요한 데이터를 제공하는 기능을 수행한다. 이후 해시값을 사용하여 경량 머클트리를 구성한다. 경량 머클트리(light-weight Merkle tree)는 두 버젼의 영상프레임 인덱스의 해당 영상프레임 변경사항을 개별 해시값의 비교 없이도 빠르게 무결성 위반을 검출할 수 있다. 이를 위해 네트워크 대역폭 효율성을 갖도록 디지털 콘텐츠의 해시값을 생성하여 경량 머클트리를 구성하고 이를 무결성 검증의 증명(proof) 수행 결과로 제시한다.

II. 관련연구

본 절에서는 디지털 콘텐츠의 무결성 검증을 위해 사용한 데이터 형식을 YUV 영상의 개요 및 특징을 설명하고 YUV 로 부터 획득된 특징의 해시값을 얻기 위한 머클트리에 대해 기술한다.

2.1 YUV 영상

YUV 영상은 색상을 나타내기 위해 삼원색을 표현하는 BGR 방식과 달리 빛의 밝기를 나타내는 휘도(Y)와 색상 신호(U, V)로 표현하는 방식이다[6]. 흑백만을 표현할 때도 RGB 는 모든 색의 데이터가 필요하기 때문에 상대적으로 많은 저장공간이 필요하다. 그러나 YUV 형식을 사용하면 자료의 크기는 1/2 정도 축소되는 장점을 갖는다. YUV 형식의 데이터는 카메라에서 얻은 RGB 데이터를 변환하여 얻고, YUV420 파일의 정보를 표시하는 헤더는 존재하지 않는다. Table 1 은 640 x 480 영상의 BGR 과 YUV 420 형식의 특징을 비교한 것이다.

Table 1. Comparison YUV420 with BGR
표.1 YUV420 과 BGR 형식 비교

	BGR	**YUV420**
width	640	640
height	480	720
channels	3	2
elements	921,600	460,800

2.2 해시함수

암호화 해시 함수(cryptographic hash function)은 해시 함수의 일종으로, 해시값으로부터 원래의 입력 값과의 관계를 찾기 어려운 다음과 같은 성질을 가진다[7][8].

- 제 1 역상 저항성(first preimage resistance): 주어진 해시 값에 대해, 그 해시 값을 생성하는 입력값을 찾는 것이 계산상 어렵다. 이 성질은 일방향함수와 연관된다.
- 제 2 역상 저항성(second preimage resistance): 입력 값에 대해, 그 입력의 해시 값을 바꾸지 않으면서 입력 변경이 계산상 어렵고, 제 2 역상 공격에 대해 안전해야 한다.
- 충돌 저항성(collision resistance): 해시 충돌에 대해 안전해야 한다. 같은 해시 값을 생성하는 두 개의 입력값을 찾는 것이 계산상 어려워야 한다.

즉, 입력값과 해시 값에 대해서, 해시 값을 변경하지 않으면서 입력값을 수정하는 공격에 대해 안전해야 한다. 이러한 성질을 가지는 해시 값은 원래 입력값을 의도적으로 손상시키지 않았는지에 대한 검증 장치로 사용할 수 있다. 제 2 역상 공격은 제 1 역상 공격에서 원본 메시지까지 주어져 있는 경우이다. 충돌 공격은 역상 공격과는 달리 해시 함수의 출력값이 고정되어 있지 않고, 해시 충돌이 일어나는 두 입력값을 찾는 공격이다. 역상 공격은 충돌 공격보다 더 어렵고, 해시 값이 같다면 입력도 같다고 할 수 있다.

따라서 보안 해시 알고리즘인 SHA256 을 사용하여 64 자리 문자열을 반환하고, 어떤 길이의 값을 입력하더라도 256 비트의 고정된 결과값을 출력하는 방식을 제안한다. 이 제안 방식에서는 입력값이 조금만 변경되어도 출력값이 완전히 달라지기 때문에 출력값을 토대로 입력값을 유추하는 것은 거의 불가능하다.

2.3 머클트리

머클트리(Merkle tree)는 블록 자료가 기록된 이후로 변경 또는 손상되지 않았음을 보장하기 위한 기법으로 블록체인에서도 활용한다[9]. 머클트리의 단말 노드는 데이터로 구성되고, 상위 노드는 자식 노드의 해시값을 갖는 자료구조이다. Figure 2 는 데이터 변경을 갖는 머클트리의 예를 보인다. 자료 K 의 해시값의 위변조가 의심되어 위변조 여부를 조사하려 할 때 필요한 정보는 4 개의 해시값(H_L, H_IJ, H_ABCDEFGH) 과 머클 루트를 사용한다.

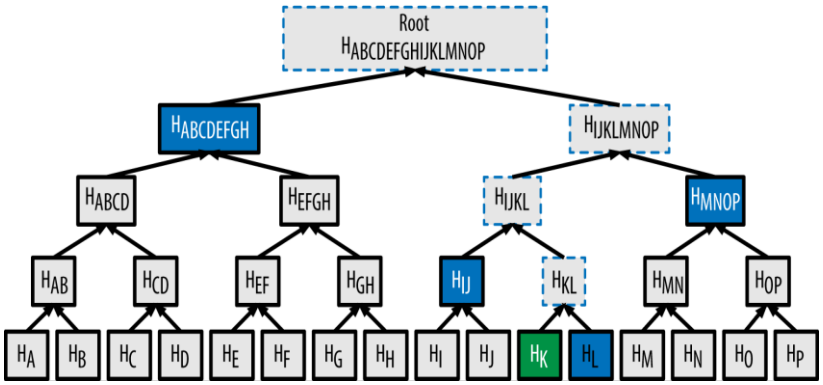

Figure 2. Example of a Merkle tree
그림 2. 머클트리 예

III. 머클트리 무결성 검사

3.1 개요

무결성 검사를 위한 Logger 와 Verifier 는 분산처리를 통해 영상 데이터의 무결성 검증을 수행한다. 검증을 위한 자료는 서버에 유지한다. 영상의 진본에 대한 확인을 위해 PKI (Public Key Infrastructure) 기반의 서명을 이용한 진본 검증을 사용한다. Logger 는 Webcam 을 통해 캡처된 영상 데이터를 YUV420 형식으로 변환, Y 프레임만 추출하여 영상에 대한 feature 를 추출한다. 이후 특징 벡터에 대해 해시값을 생성하며, 해시값의 비교를 통해 영상 프레임의 무결성을 증명하는 데 사용한다. Figure 3 은 YUV420 형식의 비디오 프레임과 Y 프레임을 보인 것이다. Figure 4 는 Figure 3 영상의 Y 프레임으로부터 생성된 feature 를 보인 것으로 feature 의 추출을 위해 canny edge 검출 기법을 사용한다.

(a) YUV420 (b) Y format r

Figure 3. Example of a Video frame with YUV420 format (a) and Y format (b)
그림 3. YUV320 영상(a) 와 Y 포멧(b) 예

Figure 4. Example of feature generated from the Y-frame
그림 4. Y 프레임으로 부터 생성된 특징 예시

Logger 는 개인키(SK, secret key)과 공개키(PK, public key)를 생성한 후 개인키는 안전한 공간에 유지하며, 공개키는 서버에 전달한다. 해당 공개키는 Verifier 에 의해 사용된다. 자료의 진본 확인을 위해 개인키과 공개키 쌍이 사용되므로 이들 키는 자료의 검증이 완료될 때까지 유지되어야 한다. Logger 는 영상 프레임 별로 해시값을 생성한 후 서명하여 서버에 저장한다. Verifier 는 해시검증을 진행하며, 추후 영상변경에 대한 검증을 위해 Merkle 트리를 구성하며, Merkle 트리 검증을 진행한다. Figure 5 은 PKI 기반 서명/검증의 처리 흐름을 보인 것이다.

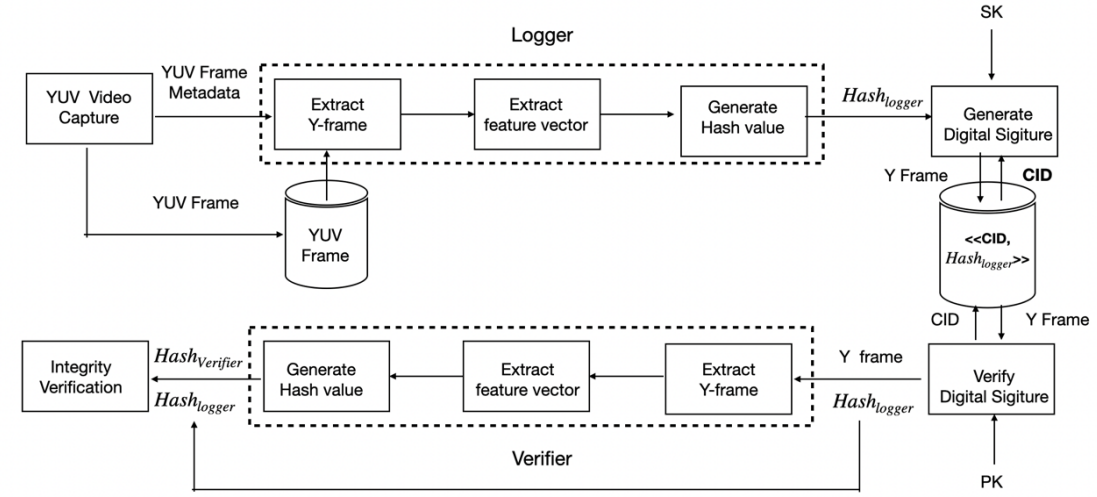

Figure 5. Overall process of logger and verifier based on PKI
그림 5. PKI 기반 logger 와 verifier 의 전체 과정

설계된 데이터 검증 처리 흐름에서 디지털콘텐츠 생성자는 머클트리 루트 노드의 해시 값(루트 해시)을 사용하여 데이터가 변경을 검출할 수 있다. 디지털콘텐츠 생성자는 디지털콘텐츠 변경검증에 필요한 자료 제공자(prover)로서 루트해시값을 생성한 후 제공하며, 검증자(verifier)는 제공된 루트해시값을 통해 데이터 무결성 검증을 수행한다. 머클트리의 구성을 위해서는 다음의 과정을 진행한다. 머클트리 생성기는 클라이언트로서 서버에 CID 와 해시값을 요청하며, 서버는 해당 CID 와 해시값을 전달한다. 머클트리 생성기는 이전에 전달된 해시값이 반복되는 경우 노드를 병합하여 대치한다. 노드 병합 후에는 머클트리 생성 후 최종 루트해시값을 서버에 전

달한다. 서버는 머클트리 루트해시값을 DB 에 삽입하여 반영한다. 다음은 머클트리 구성 과정을 보인 것이다.

1) (머클트리 생성기) 서버에게 CID 와 해시값 요청
2) (서버) CID 와 해시값을 머클트리 생성기에 전달함
3) (머클트리 생성기) 동일한 해시값이 반복되는 경우 머클트리 노드로 병합하여 대치
4) (머클트리 생성기) 머클트리를 생성한 후 해당 머클트리 루트해시값을 서버에 결과로서 전달함
5) (서버) 전달받은 머클트리 루트해시값을 DB 의 반영

3.2 실험결과 및 분석

Table 2 는 구성된 시스템의 구성 및 실험 환경을 기술한 것이다. 소켓을 이용한 TCP/IP 통신을 위해 라즈베리 파이를 통한 리눅스 환경에서 진행한다. Logger 로부터 전달된 영상정보의 무결성 검증을 위한 특징 메타정보 Maria DBMS 를 사용한다.

Table. 2 Experimental setup
표 2. 실험 환경 설정

Item		Description
S/W	OS	Raspbian GNU/Linux 11 (bullseye)
	V4L2	Standard Interface for video data
	OpenSSL 1.1.1.q	Cryptography Library
	OpenCV 4.5.1	Video Capture library
	Maria DBMS	Repository for video frame
H/W	Device	Raspberry Pi 4 Model B / 8GB RAM
	WebCAM	Logitech Webcam c270

기존의 무결성 검증을 위해 사용되는 머클트리는 주어진 n 개의 검증대상 자료에 대해 O(log n) 크기의 머클트리로 생성되어 사용된다. 이는 n 이 커짐에 따라 데이터 검증을 위한 검증 자료의 전달에 대한 네트워크 대역폭 제약을 발생시킨다. 이 논문에서는 이를 해결하기 위해 이진 머클트리를 대신하여 연속적인 해시값에 대한 중복을 제거한 후 머클트리를 구성한 후 사용한다. Figure 6 은 주어진 영상 프레임으로부터 생성된 머클트리를 보인 것이다. Figure 7 은 머클트리는 중복을 제거한 후 구성된 경량 머클트리를 보인 것이다. Figure 8 은 머클트리의 루트 해시값을 사용하여 검증을 수행한 후 무결성과 진본확인이 이루어진 후 처리 결과를 보인 것이다.

Figure 6. Example of Merkle Tree
그림 6. 머클트리 예시

Figure 7. Light-weight Merkle Tree
그림 7. 경량 머클트리 예시

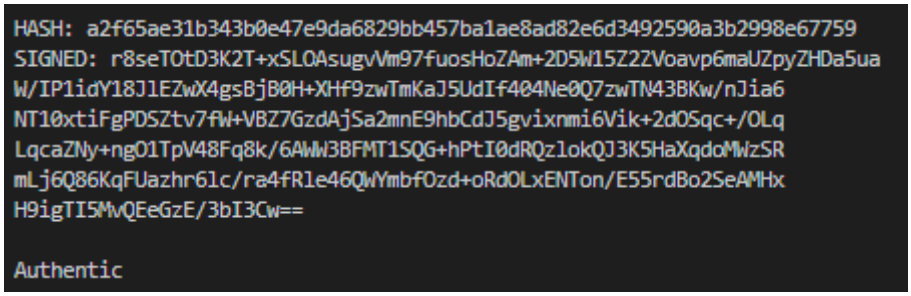

Figure 8. Result of verification
그림 8. 검증 결과

IV. 결 론

데이터 무결성은 해당 디지털 콘텐츠의 정확성과 일관성을 유지하고 보증하는 것으로 데이터 무결성의 유지를 위해서는 허가 받지 않은 사용자에 의해 데이터가 수정되었는지를 감지하는 기능이 필요함을 제시하였다. 디지털 콘텐츠의 활용을 위해서는 해당 데이터 무결성 검증은 필수적이며, 검증 자료의 네트워크 대역폭 효율성 보장이 필요하다. 이를 위해 무결성 검증기능을 제안하였으며, logger, verifier 와 서버로 구성된 시스템을 설계하고 구현하였다. 무결성 검증을 위한 작업처리 부하를 줄이기 위해 프레임별로 구성된 Merkle 트리를 활용하여 검증 소요시간을 최소화하였으며, 검증의 효율성을 위해 프레임별로 구성된 Merkel 트리를 활용하여 무결성 위반을 효율적으로 발견하였다.

V. Acknowledgement.

본 논문은 중소벤처기업부(중소기업기술정보진흥원, S3252344) 2022 년도 산학연 Collabo R&D 사업의 산업현장 영상데이터의 증거능력 확보를 위한 무결성 및 진본 검증 솔루션 개발과제의 지원을 받아 수행된 연구임

VI. 참고문헌

[1] ″Digital Transformation of the Future of Work, 2019,″ KEPCO Journal on Electric Power and Energy, vol. 6, no. 1, pp. 1−5, Mar. 2020.

[2] Alfred Zimmermann, Rainer Schmidt, Lakhmi C. Jain, Architecting the Digital Transformation - Digital Business, Technology, Decision Support, Management. Intelligent Systems Reference Library 188, ISBN 978-3-030-49639-5, Springer, 2021.

[3] S. Milani, M. Fontani, P. Bestagini, M. Barni, A. Piva, M. Tagliasacchi, and S. Tubaro, “An overview on video forensics,” APSIPA Transactions on Signal and Information Processing, vol. 1, 2012.
[4] A. Gironi, M. Fontani, T. Bianchi, A. Piva, and M. Barni, “A video forensic technique for detecting frame deletion and insertion,” in IEEE 2014 International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, pp. 6226–6230, 2014.
[5] T.Y.Kim, J.I.Hong, M.G.Kang, S.H.Song, J.H.Lee and S.T.Kim, “Integrity Support System for Blockchain-based explainable CCTV Video,” The Journal of The Institute of Internet, Broadcasting and Communication, vol. 21, no. 3, pp. 15–21, Jun. 2021.
[6] Toda, M., Tsukada, M., Inoue, A. & Suzuki, T. ,High dynamic range rendering for YUV images with a constraint on perceptual chroma preservation, IEEE. ICIP., pp. 1817-1820, ISBN: 978-1-4244-5654-3, 2009.
[7] L.V. Cherckesova, O.A. Safaryan, N.G. Lyashenko, D.A. Korochentsev, Developing a New Collision-Resistant Hashing Algorithm. Mathematics 10, 2769, 2022. https://doi.org/10.3390/math10152769.
[8] P. Rogaway, T. Shrimpton, "Cryptographic Hash-Function Basics: Definitions, Implications, and Separations for Preimage Resistance, Second-Preimage Resistance, and Collision Resistance", 2004.
[9] R. C. Merkle, “A digital signature based on a conventional encryption function,” in Advances in Cryptology - CRYPTO’87, C. Pomerance, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 369–378, 1988.

저자소개

강윤희 (*Yun-Hee Kang*)

1993년 8월 동국대학교 대학원 컴퓨터공학과 석사
2002년 8월 고려대학교 대학원 컴퓨터과학과 박사
2000년 3월~현재 백석대학교 컴퓨터공학부 교수

관심분야 : 분산시스템, 기계학습, 블록체인

장은영 (*Eun-Young CHANG*)

1991년 10월 ~ 현재 : 국립공주대학교 전기전자제어공학부 교수
1993년 2월 : 한국항공대학교 항공전자과(공학박사)

관심분야 : ICT 융합플랫폼, 이동통신, 무선통신시스템

권태언 (*Taeun Kwonk*)

2003 년 2 월 : 카톨릭 관동대학교 전자공학전공 (학사)
2022 년 1 월 ~ 현재 : ㈜하스퍼 기업부설연구소 수석연구원

관심분야 : Image Processing, machine learning Algrorithm