신뢰성 영상자료를 위한 어노테이션 기법

[1*]강윤희, [2]권태언

Annotation Method for Reliable Video Data

[*1] Yun-Hee Kang, [2] Taeun Kwon

요 약

최근 인공지능 활용 증가로 조직 내부의 AI TRiSM 보장 데이터 관리가 중요해지고 있으며, 이에 따라 데이터 신뢰성 확보는 데이터 기반 의사결정의 필수 요구사항으로 등장하였다. 생성된 디지털 콘텐츠는 신뢰성을 갖지 않는 인터넷을 통해 디지털 콘텐츠 저장소가 위치한 클라우드에 전송되어 다양하게 활용된다. 그러나 기존의 디지털 콘텐츠 시스템은 자료훼손에 따른 내용 수정을 확인하는 데이터 이상감지 기능을 제공하기 쉽지 않다. 이 논문에서는 데이터 어노테이션의 기능 확장을 통해 영상데이터의 신뢰성을 보증하기 위한 기법을 설계한다. 설계된 어노테이션 기법은 webUI 방식으로 gRPC 기반 요청 및 응답을 처리할 수 있도록 프로토타입을 구성하여 주어진 영상의 분류 레이블 및 머클트리를 생성한다.

Abstract

With the recent increase in the use of artificial intelligence, AI TRiSM data management within organizations has become important, and thus securing data reliability has emerged as an essential requirement for data-based decision-making. Digital content is transmitted through the unreliable Internet to the cloud where the digital content storage is located, then used in various applications. When detecting anomaly of data, it is difficult to provide a function to check content modification due to its damage in digital content systems. In this paper, we design a technique to guarantee the reliability of video data by expanding the function of data annotation. The designed annotation technique constitutes a prototype based on gRPC to handle a request and a response in a webUI that generates classification label and Merkle tree of given video data.

Keywords: *Data integrity, Data annotation, AI TRiSM, gRPC, Merkle tree*

[1*교신저자] 백석대학교 컴퓨터공학부 교수(yhkang@bu.ac.kr)
[2](주)하스퍼(peterkwon@harsper.co.kr)

Received: Dec. 12, 2023, Revised: Dec. 19, 2023, Accepted: Jan. 02, 2024

I. 서론

AI TRiSM 은 Artificial Intelligence(AI) Trust, Risk, and Security Management 의 약자로, AI 모델 거버넌스, 신뢰성, 공정성, 효율성, 개인 정보보호, 데이터 보호 및 신뢰성을 지원하고 가능하게 하는 정책 및 프레임워크이다. 최근 인공지능 활용 증가로 조직 내부의 AI TRiSM 보장 데이터 관리가 중요해지고 있으며, 이에 따라 데이터 신뢰성 확보는 데이터 기반 의사결정의 필수 요구사항으로 등장하였다[1]. 이러한 요구사항의 만족을 위해서는 활용 데이터의 무결성과 데이터 생산 과정에서의 진본 확인은 필수적이다[2][3][4].

AI TRiSM 보장을 위해서는 수집된 데이터셋에 대한 데이터 이상 감지(Data Anomaly Detection)을 지원하여야 한다[1]. 데이터 이상 감지는 데이터 변경에 따른 문제를 감지하고 식별하는 기능으로 AI 실무자가 데이터의 전체적인 문제를 파악하고 효과적인 결정을 내릴 수 있도록 한다. 그러나 기존의 디지털 콘텐츠 시스템은 자료훼손에 따른 내용 수정을 확인하는 데이터 이상감지 기능을 제공하기 쉽지 않다. 생성된 디지털 데이터는 일반적으로 신뢰성을 갖지 않는 인터넷을 통해 데이터 저장소가 위치한 클라우드에 전송되어 다양하게 활용된다.

데이터 어노테이션(Data annotation)은 학습 데이터의 품질을 보증하기 위한 목적으로 사용되지만, AI TRiSM 관점에서 데이터 이상 감지를 위한 데이터 무결성 검증 검사 정보를 추가할 수 있다. 주어진 메타데이터에 데이터 변경 관련 무결성 검사 정보를 추가하는 것은 다음의 이점을 갖는다.

- 외부의 데이터 접근에 대한 보안 기능의 강화
- 데이터 훼손에 대한 실시간 모니터링과 위험 검출
- 데이터기반 의사결정 시스템 내의 취약성 저감

이 논문에서는 데이터 어노테이션의 기능 확장을 통해 영상 데이터의 신뢰성을 보증하기 위한 기법을 설계한다. 설계된 어노테이션 기법은 webUI 방식으로 gRPC 기반 요청 및 응답을 처리할 수 있도록 프로토타입을 구성하여 주어진 영상의 분류 레이블 및 머클트리(Merkle tree)를 생성한다. 구성된 프로토타입은 네트워크 대역폭 효율성을 갖도록 디지털 콘텐츠의 해시값을 생성하여 머클트리를 구성하고, 이를 데이터 이상감지 결과로 제시한다.

II. 관련연구

본 절에서는 디지털콘텐츠 분류를 위한 전이학습(Transfer Learning), 무결성 검증을 위한 머클트리 및 설계된 어노테이션 기법의 프로토타입 개발 환경인 gRPC 을 기술한다.

2.1 전이학습

전이학습은 기학습된 신경망 네트워크의 일부를 상이한 영역에 적용하여 모델을 학습하는 딥러닝 학습기법 중의 하나이다 [5]. 해당 기법은 학습 데이터의 수가 적거나 혹은 불균형인 경우 모델을 학습하는데 유용하고, 미세한 조정을 통해 적은 학습 횟수로도 학습 효율을 향상할 수 있는 이점이 있다. 본 연구는 전이학습을 적용하기 위해, 사전에 개발된 중추 네트워크 모델(Backbone Network Model)을 사용한다[6].

이 논문에서 적용된 전이학습 모델은 ResNet 로서 학습 모델의 네트워크가 깊어질수록 기울기 소멸(Gradient Vanishing) 현상으로 인하여 학습률 및 성능이 감소되는 제약점을 해결하기 위해 고안되었다. ResNet 은 학습모델과 입력 값의 차이인 잔차 (Residual)가 0 이 되도록 학습을 수행하는 모델로서 Skip connection 연산을 통해 컨볼류션을 위해 배치된 앞단 레이어의 출력 값 F(X)에 피드 포워드(feed forward) 한 값인 X 를 더해준다. Figure 1 은 ResNet 의 개념적 구조를 보인 것이다.

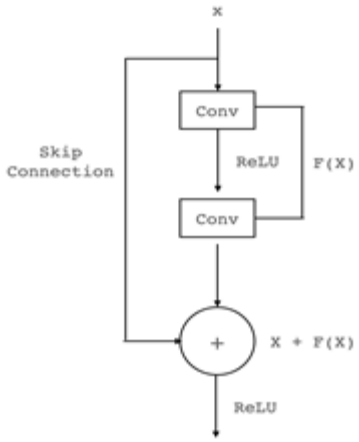

Figure 1. Conceptual structure of ResNet
그림 1. 개념적인 ResNet 구조

2.2 머클트리

머클트리는 자식 노드들이 해시값을 갖는 트리 형태의 자료구조이다. 머클트리의 단말 노드는 데이터로 구성되고, 상위 노드는 자식 노드의 해시값을 갖는 자료구조이다. 이진 머클트리는 두 버젼의 영상프레임 인덱스의 해당 영상프레임 변경사항을 개별해시값의 비교없이도 빠르게 검출할 수 있다. 구성된 머클트리의 최상위에 위치한 최종 값을 루트, 하부의 값들을 리프라고 한다. 데이터 변경을 갖는 머클트리는 해시값의 위변조가 의심되는 경우 이를 검증하는 과정에서 머클 트리루트를 참고하여 머클 경로(Merkle path)를 통해 위변조 여부를 조사할 수 있다. 머클트리는 블록 자료가 기록된 이후로 변경 또는 손상되지 않았음을 보장하기 위한 기법으로 블록체인에서도 활용한다[7].

Figure 2 는 k0 에서 k4 의 5 개 자료를 갖는 머클트리 구성 과정을 보인 것이다. 암호화 해시 함수(Cryptographic hash function)은 해시 함수의 일종으로, 해시 값으로 부터 원래의 입력값과의 관계를 유추하기 어려운 성질을 가지며, 해시 값을 변경하지 않으면서 입력값을 수정하는 공격에 대해 안전하다[8]. 머클트리 구성을 위한 해시함수 h 는 암호화 해시 알고리즘인 SHA256 을 사용하여 고정된 길이의 결과값을 출력하도록 한다. SHA256 은 입력값의 변경 시 해당 해시값이 변경되기 때문에 해시함수의 출력값을 기반으로 입력값을 유추하는 것은 거의 불가능하다.

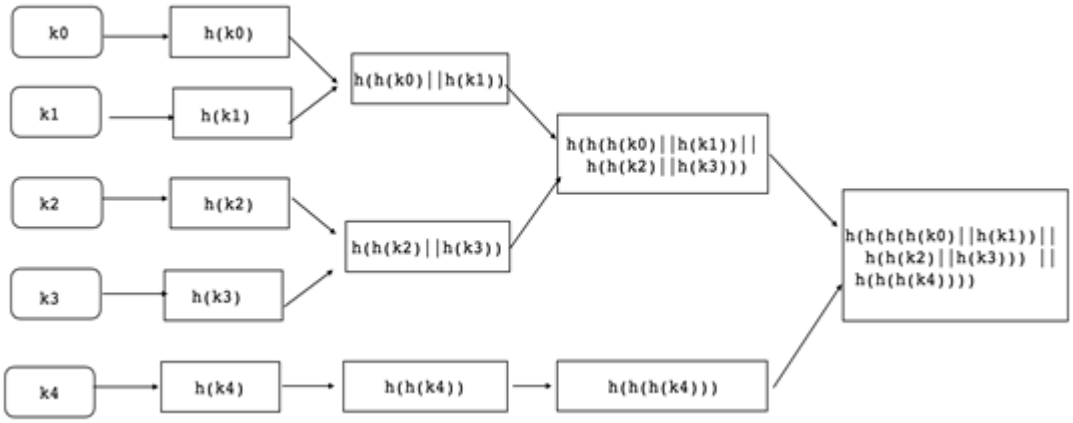

Figure 2. Process for building a Merkle tree
그림 2. 머클트리 구성 과정

2.3 gRPC

gRPC 는 Google 에서 개발한 오픈소스 원격 프로시저 호출인 RPC(Remote Procedure Call) 프레임워크로 다양한 분야의 서비스 지향 설계를 제공한다 [9][10]. Figure 3 은 gRPC 응용 구조를 보인 것으로 요청을 처리하는 서비스(Service)는 gRPC 서버에서 동작하며, 클라이언트에서는 gRPC Stub 이 서비스 요청을 위해 동작한다. Stub 은 gRPC 에서 클라이언트와 서버 간의 통신을 추상화하고 단순화하는 데 사용되는 코드 조각으로 클라이언트와 서버 모두에 위치하며, 호출 및 반환 과정을 처리하는 역할을 수행한다.

gRPC 에서는 HTTP/2 를 사용하여 요청 및 응답처리에서 경량 네트워크 공간 및 압축의 이점을 제공한다. PB(Protocol Buffer)는 RPC 서비스를 정의하고 gRPC 클라이언트 인터페이스를 생성하며, 다양한 언어와 환경 간의 바이너리 통신의 제약을 위한 상호운영성을 제공하다. PB 는 IDL(Interface Definition Language)로서 사용한다 [11]. PB 를 사용한 서비스 정의는 매개 변수 및 반환 유형을 기술하고 원격으로 호출할 수 있는 메서드를 지정한다. 다음은 gRPC 의 주요한 특징이다.

- gRPC 는 원격에 있는 애플리케이션의 메서드를 로컬 메서드인 것처럼 직접 호출할 수 있으므로 분산 애플리케이션 및 서비스를 쉽게 작성할 수 있다.
- gPRC 클라이언트와 서버는 클라우드 환경에서 데스크탑, 모바일까지 다양한 환경에서 실행할 수 있고 다양한 언어를 지원한다.
- 데이터를 직렬화 기법을 제공하는 PB 는 마샬링(marshaling) 속도 및 코드 크기 측면에서 텍스트 기반인 JSON 보다 효율적이다.

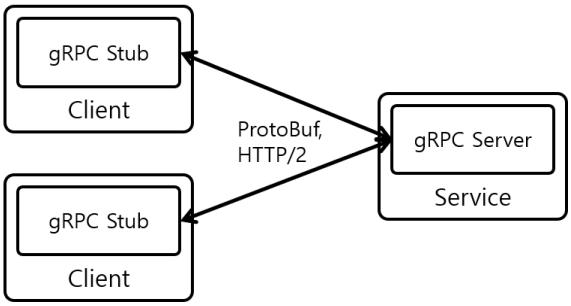

Figure 3. Example of gRPC application
그림 3. gRPC 응용 예

III. 설계된 신뢰성을 갖는 영상 데이터 어노테이션 기법

3.1 개요

영상 데이터의 처리 시스템에서 Logger 는 영상프레임별로 해시값을 생성한 후 서명하여 Verifier 에 전달한다. Verifier 는 해시검증을 진행하며, 추후 영상변경에 대한 검증을 위해 머클트리를 구성한다. 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 증명자(Prover)로서 루트해시값을 생성한 후 제공하며, 검증자(Verifier)는 제공된 루트해시값을 통해 데이터 무결성 검증을 수행한다.

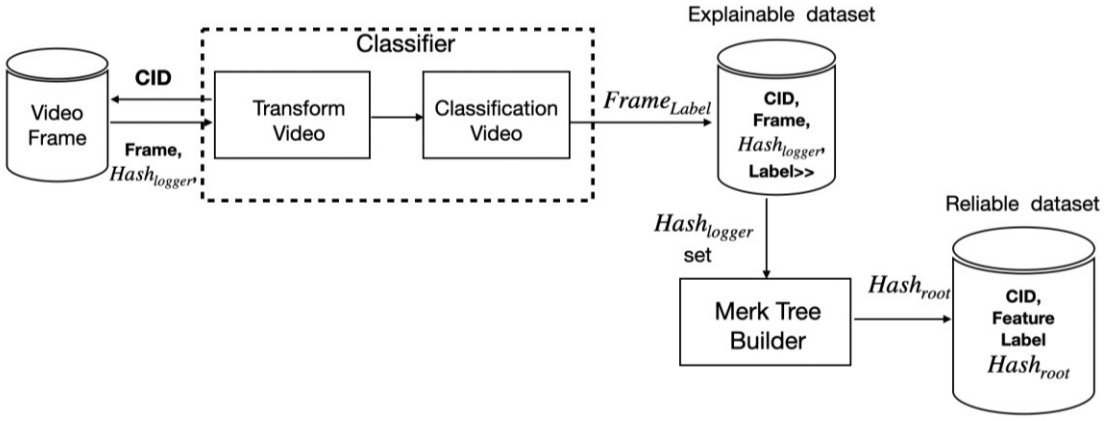

Figure 4. Process of designed annotation
그림 4. 설계된 어노테이션 프로세스

Figure 4 는 Verifier 에서 수집된 영상 데이터의 분류 레이블과 머클트리 루트를 생성하여 무결성 검사를 수행하는 설계된 어노테이션 프로세스를 보인 것이다. 어노테이션 프로세스에서 영상의 분류 및 개별 해시값은 어노테이터(Annotator)를 통해 머클트리 루트가 생성되어 신뢰가능한 데이터셋에 유지된다. 어노테이터는 영상 분류를 위한 분류기(classifier)와 머클트리 생성기(Merkle tree builder)로 구성된다. 어노테이션 프로세스를 통해 설명가능한 데이터셋(Explainable dataset)에 추가로 머클트리 루트가 포함된 신뢰가능한 데이터셋(Reliable dataset)을 생성한다.

3.2 실험결과

설계된 신뢰성 보장을 위한 영상 어노테이션 기법은 Logger 로부터 획득된 영상데이터를 대상으로 검증을 수행하는 모듈인 Verifier 에 데이터 어노테이션 기법을 추가하여 실험을 수행한다. 이과정에서 영상 진본 확인을 위해서는 PKI (Public Key Infrastructure) 기반의 서명을 이용한 진본검증을 사용한다. PKI 는 공개 키 알고리즘을 통한 암호화 및 정보 통신에 필요한 Logger 와 Verifier 간의 인증을 제공하기 위한 보안 환경을 제공한다. 이후 영상 자료의 분류 결과 및 머클트리 루트값은 메타데이터로 추가된다. 설계된 기법은 webUI 방식으로 gRPC 기반 요청 및 응답을 처리할 수 있도록 프로토타입을 구성하여 영상 분류 및 머클트리 구성한다.

영상분류와 머클트리 서비스는 Python3.8 으로 작성되며, gRPC 통신을 통해 수행한다. 영상분류와 머클트리 서비스의 webUI 연동을 위해 API 서버를 구성하여 REST(Representational State Transfer) 요청을 처리하도록 한다. HTTP 기반의 REST 요청의 처리는 API 서버를 통해 전달되도록 한다. API 서버 구성을 위해서는 python 기반의 웹응용 개발 프레임워크인 fastAPI 를 사용하여 개발한다.

프로토타입의 물리적 구성은 에지 장치인 리눅스를 운영체제로 갖는 Jetson Orin Nano 를 사용한다. Logger 로 부터 전달된 영상정보의 무결성 검증을 위한 특징 메타정보의 유지를 위해 Maria DBMS 를 사용한다.

Figure 5 은 gRPC 을 위한 PB 서비스 명세를 보인 것으로 영상분류를 위한 rpc 서비스인 inference 는 요청 InferenceRequest 과 응답 InferenceReply 을 사용한다. InferenceRequest 은 영상 프레임은 바이트 배열인 bytes 을 자료형으로 InferenceReply 은 영상 분류 레이블 번호를 표현하기 위해 정수형 자료형 uint32 을 사용한다.

```
// The inference service definition.
service InferenceServer {
  // Sends a inference reply
  rpc inference (InferenceRequest) returns (InferenceReply) {}
}
// The request message containing the images.
message InferenceRequest {
    repeated bytes image = 1;
}
// The response message containing the classes ids
message InferenceReply {
    repeated uint32 pred = 1;
}
```

Figure 5. service description for classification service
그림 5. 분류 서비스를 위한 서비스 기술

Figure 6 은 머클트리 구성을 위한 gRPC 서비스 명세를 보인 것으로 머클트리 구성을 위한 서비스 build_mt 는 요청 MT_Request 와 응답 MT_Response 을 사용한다. MT_Request 은 영상자료의 sha256 해시값의 집합인 fingerprint 으로 구성된 문자열을 자료형으로 MT_Response 은 머클트리 루트 root 를 표현하기 위해 문자열 자료형을 사용한다.

```
service MerkleTree{
    rpc build_mt(MT_Request) returns (MT_Response);
}

message MT_Response{
  string root = 1;
}

message MT_Request {
  string fingerprint = 1;
}
```

Figure 6. service description for building Merkle tree service
그림 6. 머클트리 구성을 위한 서비스 기술

Figure 7 은 클라이언트에서 주어진 영상 데이터의 영상 분류 및 머클트리 루트의 구성을 보인 것이다. webUI 는 영상 파일선택 후 해당 영상의 분류가 수행되며, 주어진 분류영상의 sha256 해시값을 입력으로 머클트리를 생성한 후 머클트리 루트 결과를 출력한다. 영상분류와 머클트리 루트 자료는 영상에 대한 추가 정보로서 영상 메타데이터에 추가되어 저장된다. 추가된 영상 메타데이터는 주어진 영상의 설명 및 내용변경발생에 따른 무결성 검사를 통한 이상 검출에 적용한다.

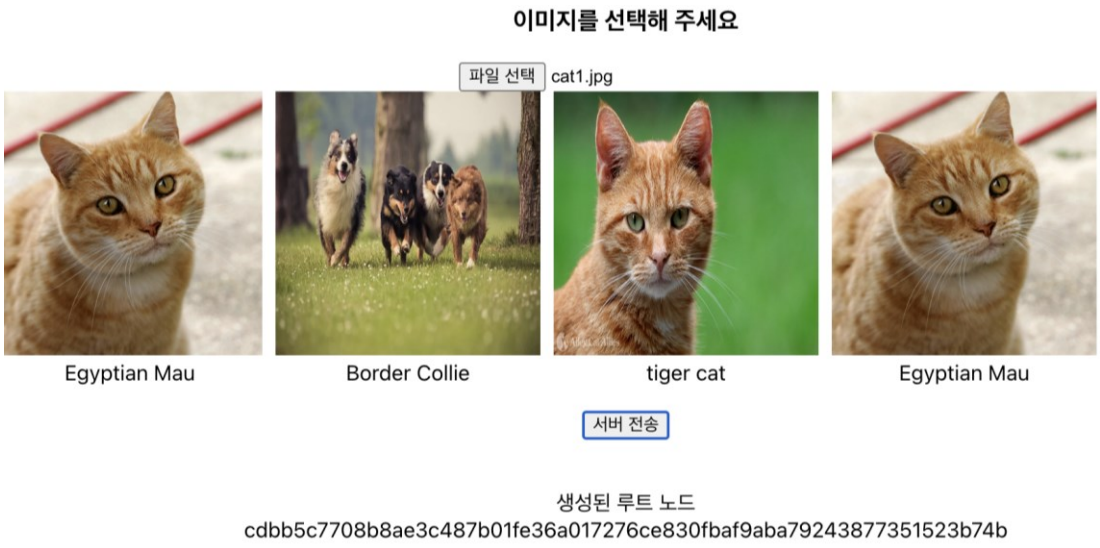

Figure 7. Example of classification label and Merkle tree root
그림 7. 분류 정보 및 머클트리 루트 구성 예시

IV. 결 론

AI TRiSM 을 지원하기 위해 데이터 무결성은 해당 디지털 콘텐츠의 정확성과 일관성이 유지되었음을 보증하기 위해 주요하며, 데이터 운영 과정에서 데이터가 수정되었는지를 감지하는 기능이 필요하다. 이를 위해 설계된 어노테이이션 기법은 영상의 분류정보를 추가하고 해당 영상에 대해 머클트리 루트값을 계산한다. 설계된 데이터 어노테이션 기법은 클라이언트-서버 기반으로 프로토타입을 구성하였으며 프로토타입을 통해 클라이언트에서 주어진 영상의 분류정보 및 머클트리 루트를 생성하였으며, 생성된 정보는 영상의 설명 및 내용변경발생에 따른 무결성 검사에 적용한다.

V. Acknowledgement.

본 논문은 중소벤처중소벤처기업부(중소기업기술정보진흥원, RS-2023-00225234) 2023 년도 산학연 CollaboR&D 사업의 산업현장의 디지털 영상데이터의 AI 기반 무결성 및 검증 기술 개발과제의 지원을 받아 수행된 연구임

VI. 참고문헌

[1] Adib Habbal, Mohamed Khalif Ali, Mustafa Ali Abuzaraida, Artificial Intelligence Trust, Risk and Security Management (AI TRiSM): Frameworks, applications, challenges and future research directions, Expert Systems with Applications, Volume 240, 2024, 122442, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2023.122442.
[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] Al-Stouhi, S., & Reddy, C. K. (2016). Transfer learning for class imbalance problems with inadequate data. Knowledge and information systems, 48(1), 201-228.
[6] Agarwal, N., Sondhi, A., Chopra, K., & Singh, G. (2021). Transfer learning: Survey and classification. In Smart Innovations in Communication and Computational Sciences (pp. 145-155). Springer, Singapore.
[7] 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.
[8] P. Rogaway, T. Shrimpton, "Cryptographic Hash-Function Basics: Definitions, Implications, and Separations for Preimage Resistance, Second-Preimage Resistance, and Collision Resistance", 2004.
[9] Brian N. Bershad, Thomas E. Anderson, Edward D. Lazowska, and Henry M. Levy. Lightweight Remote Procedure Call. ACM Trans. Comput. Syst., 8(1):37–55, February 1990.
[10] gRPC. https://grpc.io/, 2022.
[11] ingwei Wang, Hong Zhao, and Jiakeng Zhu. 1993. GRPC: a communication cooperation mechanism in distributed systems. SIGOPS Oper. Syst. Rev. 27, 3 (July 1993), 75–86. https://doi.org/10.1145/155870.155881

저자소개

강윤희*(Yunhee Kang)*

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

관심분야 : Distributed System, Artificial Intelligence , Cloud Computing

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