

Fat Client-Based Abstraction Model of Unstructured Data for Context-Aware Service in Edge Computing Environment

Do Hyung Kim[†] · Jong Hyeok Mun^{††} · Yoo Sang Park^{††} · Jong Sun Choi^{†††} · Jae Young Choi^{††††}

ABSTRACT

With the recent advancements in the Internet of Things, context-aware system that provides customized services become important to consider. The existing context-aware systems analyze data generated around the user and abstract the context information that expresses the state of situations. However, these datasets is mostly unstructured and have difficulty in processing with simple approaches. Therefore, providing context-aware services using the datasets should be managed in simplified method. One of examples that should be considered as the unstructured datasets is a deep learning application. Processes in deep learning applications have a strong coupling in a way of abstracting dataset from the acquisition to analysis phases, it has less flexible when the target analysis model or applications are modified in functional scalability. Therefore, an abstraction model that separates the phases and process the unstructured dataset for analysis is proposed. The proposed abstraction utilizes a description name Analysis Model Description Language(AMDL) to deploy the analysis phases by each fat client is a specifically designed instance for resource-oriented tasks in edge computing environments how to handle different analysis applications and its factors using the AMDL and Fat client profiles. The experiment shows functional scalability through examples of AMDL and Fat client profiles targeting a vehicle image recognition model for vehicle access control notification service, and conducts process-by-process monitoring for collection-preprocessing-analysis of unstructured data.

Keywords : Edge Computing, Fat Client, Context Aware, Unstructured Data Abstraction, Deep Learning

에지 컴퓨팅 환경에서의 상황인지 서비스를 위한 팻 클라이언트 기반 비정형 데이터 추상화 방법

김도형[†] · 문종혁^{††} · 박유상^{††} · 최종선^{†††} · 최재영^{††††}

요약

최근 사물인터넷의 발전으로 사용자 주변 상황을 인지하여 맞춤형 서비스를 제공하는 상황인지 시스템에 대한 관심이 증가되고 있다. 기존의 상황인지 시스템은 사용자 주위에서 생성되는 데이터를 분석하여 사용자 주변 상황을 표현하는 상황 정보로 추상화하는 기술이 사용되었다. 하지만 증가하는 사용자의 서비스 요구 사항에 따라 다양한 종류의 비정형 데이터의 사용이 증가하고, 사용자 주변에서 수집되는 데이터의 양이 많아지면서 비정형 데이터의 처리와 상황인지 서비스의 제공에 어려움이 있다. 이러한 사항은 딥러닝 응용에서 비정형 구조의 입력 데이터가 많이 사용되는 데서 찾아볼 수 있다. 기존 연구에서는 에지 컴퓨팅 환경에서 다양한 딥러닝 모델을 활용해 비정형 데이터를 상황 정보로 추상화하는 연구가 진행되었으나, 수집-전처리-분석 등과 같은 추상화 과정 간의 종속성으로 인해 제한된 종류의 딥러닝 모델만이 적용 가능하기 때문에 시스템의 기능적 확장성이 고려되어야 한다. 이에 본 논문은 에지 컴퓨팅 환경에서 딥러닝 기술을 활용한 비정형 데이터 추상화 과정의 기능적 확장성을 고려한 비정형 데이터 추상화 방법을 제안한다. 제안하는 방법은 데이터 처리가 분산되어 있는 에지 컴퓨팅 환경에서 수집과 전처리 과정을 수행할 수 있는 팻 클라이언트 기술을 사용하여 추상화 과정의 수집-전처리 과정과 분석 과정을 분리하여 수행하는 것이다. 또한 분리된 추상화 과정을 관리하기 위해 수집-전처리 과정을 수행하는 데 필요한 정보를 팻 클라이언트 프로파일로 제공하고, 분석 과정에 필요한 정보를 분석 모델 설명 언어(AMDL) 프로파일로 제공한다. 두 가지 프로파일을 통해서 추상화 과정을 독립적으로 관리하여 상황인지 시스템의 기능적 확장성을 제공한다. 실험에서는 차량 출입 통제 알림 서비스를 위한 차량 이미지 인식 모델을 대상으로 팻 클라이언트 프로파일과 AMDL 프로파일의 예제를 통해 시스템의 기능적 확장성을 보이고, 비정형 데이터의 추상화 과정별 세부사항을 보인다.

키워드 : 에지 컴퓨팅, 팻 클라이언트, 상황인지, 비정형 데이터 추상화, 딥러닝

※ 이 논문은 2019년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구 사업임(No.2019R1A2C1007861).

† 비회원 : 송실대학교 컴퓨터학과 석사과정

†† 준회원 : 송실대학교 컴퓨터학과 박사과정

††† 정회원 : 송실대학교 컴퓨터학부 부교수

†††† 종신회원 : 송실대학교 컴퓨터학부 교수

Manuscript Received : December 3, 2020

Accepted : January 14, 2021

* Corresponding Author : Jong Sun Choi(jongsun.choi@ssu.ac.kr)

1. 서론

최근 사물인터넷의 발전으로 헬스케어, 스마트 팜, 스마트 시티 등 다양한 분야에서 사용자 주변 상황을 인식하여 적절한 서비스를 제공하기 위한 많은 연구들이 진행되고 있다 [1-3]. 기존의 많은 연구들은 사물인터넷 환경에서 사용자 주변의 수많은 센서로부터 생성되는 데이터를 분석하고 추론하

여 시간, 위치, 장소와 같은 사용자의 상황을 나타내는 상황 정보로 추상화하는 상황인지 시스템을 통해 사용자에게 적절한 서비스를 제공한다[4, 5].

그러나 상황인지 시스템을 통해 맞춤형 서비스를 제공하는 기존 연구들은 다음과 같은 문제로 서비스 제공에 어려움이 있다. 증가하는 사용자의 서비스 요구사항에 따라 비정형 데이터의 사용이 증가하면서 비정형 데이터의 처리에 어려움이 있다[6, 7]. 비정형 데이터를 상황 정보로 추상화하는 과정은 딥러닝(Deep-learning) 기술과 같이 데이터의 특징과 패턴 등의 학습을 통해 진행되며, 추상화 과정은 비정형 데이터의 수집, 전처리, 분석의 세 가지 과정으로 진행된다. 이러한 추상화 과정에서 사용되는 딥러닝 모델은 학습을 위해 많은 컴퓨팅 자원이 요구되며, 학습된 딥러닝 모델을 통한 분석 과정에도 많은 컴퓨팅 자원이 요구된다[8]. 또한 딥러닝 기반 비정형 데이터 추상화 과정의 수집, 전처리, 분석 과정 사이의 종속성으로 인하여 시스템의 기능적 확장성을 제공하는데 어려움이 있다. 딥러닝 기반 상황 정보 추상화 과정에 사용되는 분석 모델의 실행 환경과 입력 데이터에 따라 데이터 수집 디바이스와 전처리 과정에 많은 영향을 미치기 때문에 시스템의 기능적 확장성을 한정시킨다.

이러한 문제는 에지 컴퓨팅(Edge Computing) 환경에서 딥러닝 기술을 적용하여 해결할 수 있다. 에지 컴퓨팅은 데이터가 생성되는 네트워크 에지에서 데이터 저장 및 분석과 같은 계산 작업을 수행하는 컴퓨팅 패러다임이다[9]. 에지 컴퓨팅은 일반적으로 디바이스 계층, 에지 계층, 클라우드 계층의 3계층으로 구성되어 있으며 계층별로 작업을 분산하여 일부 계산 작업을 처리하기 때문에 딥러닝 분석과 같이 필요한 계산 능력이 많이 요구되는 작업을 처리하는데 이점을 가진다[10, 11]. 그러나 에지 컴퓨팅 환경에서 딥러닝 기술을 적용한 비정형 데이터 추상화 과정을 수행하면서 시스템의 기능적 확장성을 제공하기 위해서는 다음과 같은 요구사항이 있다. 사용자의 서비스 요구사항에 따라 추상화 과정에서 사용되는 분석 모델이 바뀌거나 업데이트되었을 때 시스템의 기능적 확장성을 제공하기 위해서는 분석 모델의 실행 환경의 설정 및 관리가 요구된다[12, 13]. 또한 분석 모델의 입력형태로 수집 데이터를 전처리 할 수 있어야 한다[14].

데이터 처리과정이 분산되어 있는 에지 컴퓨팅 환경에서 이러한 요구사항을 해결하기 위해서는 수집과 전처리 작업을 인스턴스화 하여 수행하고 관리해야 해야 하고, 이는 기존의 썬 클라이언트(Thin client)가 아닌 별도의 작업을 수행할 수 있는 팻 클라이언트(Fat client)기술이 필요하다[15].

이에 본 논문은 에지 컴퓨팅 환경에서 상황인지 시스템의 기능적 확장성을 고려한 팻 클라이언트 기반 비정형 데이터 추상화 방법을 제안한다. 제안하는 방법은 수집, 전처리 과정에 필요한 정보를 팻 클라이언트 프로파일로 제공하고, 분석 모델의 실행환경 관리에 필요한 정보를 분석 모델 설명 언어(이하 AMDL)로 제공하여 추상화 과정을 관리하는 것이다. 두

가지 프로파일을 통해 추상화 과정을 독립적으로 관리하기 때문에 상황인지 시스템의 기능적 확장성을 제공할 수 있다.

2. 관련 연구

본 절에서는 에지 컴퓨팅 환경에서 딥러닝 기술을 적용하여 상황인지 서비스를 제공하는 기존 연구들의 시스템 요구사항들을 확인하고, 시스템의 기능적 확장성 측면에서 시스템을 구성하기 위해 필요한 요구사항을 확인한다.

에지 컴퓨팅 환경에서 딥러닝 기술을 적용해 상황인지 서비스를 제공하는 시스템은 보편적으로 디바이스, 에지, 클라우드의 3계층 구조를 가지고 있으며, 비정형 데이터 수집, 전처리, 분석 과정을 통해 비정형 데이터를 상황 정보로 추상화하여 맞춤형 서비스를 제공한다. 이러한 시스템은 지능형 차량 서비스, 스마트 헬스케어 서비스 등의 분야에서 활용된다.

Salman[16], Chen[17], Hesham[18]의 연구에서는 에지 컴퓨팅 환경에서 지능형 차량 서비스를 위해 수집되는 비정형 데이터의 상황 정보 추상화 과정에 딥러닝 기술이 활용되었다. 에지 컴퓨팅 환경에서 딥러닝 기술을 활용하여 지능형 차량 서비스 제공을 위한 컴퓨팅 구조를 제공하며 이를 통해 활용도가 낮은 개별 차량의 컴퓨팅 자원을 활용하여 실시간 데이터 전달 및 이동성을 제공한다. 이러한 시스템은 차량 내에 배치되어 있는 장치들을 통해 데이터를 수집하고 분석하는 추상화 과정을 수행하기 때문에 시스템의 기능적 확장성을 제공하기 위해서는 사용자 요구사항에 따른 수집 장치의 추가 및 변경에 대한 추상화 과정의 관리가 필요하다.

헬스케어 분야 또한 에지 컴퓨팅 환경에서 상황인지 기술과 딥러닝 기술을 접목하여 서비스를 제공하는 연구를 진행하고 있다. Uddin[19]와 Min[20]의 연구는 에지 컴퓨팅 환경에서 사용자 주변 환경정보를 상황 정보로 추상화하여 사용자 맞춤형 헬스케어 서비스를 제공하기 위해 상황 정보 추상화 과정에 딥러닝 기술을 활용하였다. 또한 Uddin은 추상화 과정에 사용되는 딥러닝 모델을 위해 센서 데이터 수집 공개 데이터 세트를 제공한다. 딥러닝 기반 분석에 사용되는 센서 데이터는 제공된 공개 데이터 세트를 통해 관리되어 상황 정보 추상화의 수집 단계의 확장성을 제공한다.

에지 컴퓨팅 환경에서 딥러닝 기술을 활용한 상황인지를 통해 서비스를 제공하는 또 다른 연구로 Hosseini[21]과 Tuli[22]의 연구는 에지 컴퓨팅 환경에서 이미지 센서 디바이스로부터 수집된 데이터를 딥러닝 기반 분석 과정을 통해 상황 정보로 추상화 하고 맞춤형 서비스를 제공한다. 수집된 이미지 데이터는 일반 그래픽 형식이며 딥러닝 모델에 대한 입력으로 사용하기 위해서는 전처리 과정이 필요하다. Hosseini와 Tuli의 연구는 시스템에 적용되는 딥러닝 모델을 위한 전처리 과정을 제안하였고 이는 수집 데이터와 딥러닝 모델의 실행환경에 따라 바뀔 수 있기 때문에 서비스 요구사항에 따른 특성 및 목적이 다른 딥러닝 모델을 사용하기에 어려움이 있다.

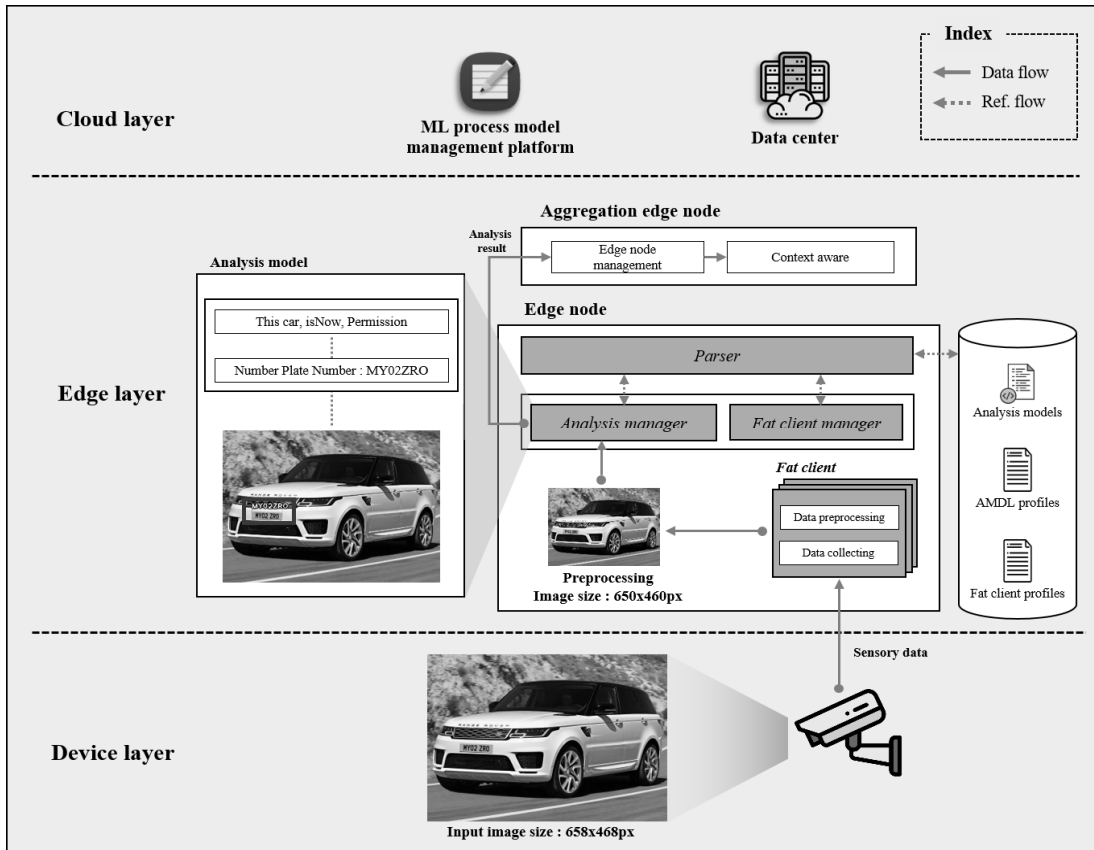


Fig. 1. Overview of Fat Client based Context-Aware System in Edge Computing

에지 컴퓨팅 환경에서 상황인지 시스템을 위해 딥러닝 기술을 활용하여 수집 데이터를 상황 정보로 추상화하는 기존 연구들은 변화하는 사용자 요구사항과 서비스 환경에 따라 추상화 과정의 규칙의 변화와 딥러닝 모델의 실행환경의 변화가 발생할 수 있다. 이는 추상화 과정 사이의 종속성으로 인해 발생하기 때문에 상황인지 시스템의 기능적 확장성을 제공하기 위해서는 추상화 과정의 관리가 필요하다.

3. 에지 컴퓨팅 환경 팻 클라이언트 기반 비정형 데이터 추상화 방법

본 절에서는 에지 컴퓨팅 환경에서 비정형 데이터 처리를 통해 상황인지 서비스를 제공하는 시스템의 요구사항을 검토하여 팻 클라이언트 기반 비정형 데이터 추상화 방법을 적용한 상황인지 시스템의 개요와 시스템 구조를 소개하고 구조의 각 계층을 구성하고 있는 모듈들에 대해 설명한다. 시스템의 각 계층별 요구사항은 다음과 같다.

- 디바이스 계층(Device layer)
 - 이기종 디바이스의 검색 및 데이터 제공
- 에지 계층(Edge layer)
 - 비정형 데이터의 수집, 전처리 및 분석을 통한 상황 정보 추상화

- 로컬 데이터 저장 및 관리
- 상황 정보 기반의 상황인지 서비스
- 팻 클라이언트 생성 및 관리
- 분석 모델 관리

- 클라우드 계층(Cloud layer)
 - 데이터 집계 및 데이터 웨어하우스
 - 분석 모델, AMDL 프로파일 생성 및 관리

3.1 시스템 개요

에지 컴퓨팅 환경 팻 클라이언트와 딥러닝 기술을 활용한 상황인지 시스템은 비정형 데이터 수집, 전처리, 분석 과정을 통해 비정형 데이터를 상황 정보로 추상화하여 상황인지 과정을 수행하며 추상화 과정에서 팻 클라이언트와 딥러닝 모델이 사용된다. 이때, 딥러닝 모델을 활용한 추상화 과정을 팻 클라이언트 프로파일과 AMDL 프로파일을 통해 관리함으로써 분석 모델 변경 및 추가에 대한 상황인지 시스템의 기능적 확장성을 제공한다. 시스템 개요는 Fig. 1과 같다.

디바이스 계층은 주변 센서, 웨어러블 디바이스, 카메라 등 이기종 디바이스로부터 상황인지를 하기 위한 비정형 데이터를 제공한다. 수집된 데이터는 유무선 통신 프로토콜을 통해 데이터를 에지 계층으로 전송한다.

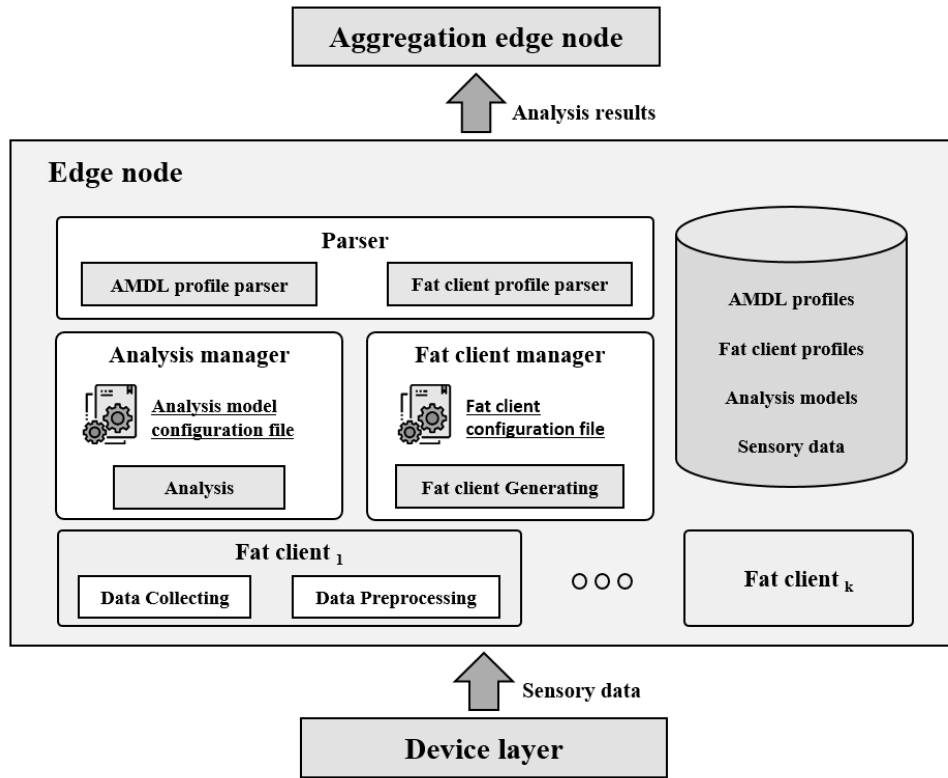


Fig. 2. System Architecture of Edge Node for Unstructured Data Abstraction

에지 계층은 팻 클라이언트 프로파일을 기반으로 각각의 비정형 데이터의 수집과 전처리 과정을 수행하는 팻 클라이언트 인스턴스를 생성하여 수집과 전처리 과정을 수행한다. AMDL 프로파일을 통해 분석 모델을 사용할 수 있는 환경을 만들고, 팻 클라이언트에서 전처리한 비정형 데이터를 상황 인지 서비스를 제공하기 위한 상황 정보로 추상화하는 역할을 한다. 에지 계층의 통합 에지 노드는 비정형 데이터의 분석을 통해 추상화된 상황 정보를 바탕으로 상황을 인지하고 적절한 서비스를 제공한다.

클라우드 계층은 수집된 모든 데이터, 분석한 정보, 상황 정보 등 모든 데이터의 통합 저장 및 관리를 수행한다. 또한 수집된 데이터를 바탕으로 분석 모델과 AMDL 프로파일의 생성 및 갱신 등의 관리를 하고 에지 계층에 배포하는 역할을 한다.

3.2 시스템 구성

제안하는 시스템은 에지 계층에서 에지 노드(Edge Node)와 통합 에지 노드(Aggregation Edge Node) 2개의 영역으로 나뉘지며, 각 영역의 구성요소들에 대한 상세한 내용은 다음과 같다.

1) 에지 노드

에지 노드는 비정형 데이터의 수집, 전처리, 분석의 추상화 과정의 수행 및 관리를 위해 필요로 구성되어 있으며 로컬 데이터의 저장 및 관리를 위한 저장소를 가진다. 에

지 노드의 자세한 구조는 Fig. 2와 같다. 구성요소는 팻 클라이언트(Fat client), 파서(Parser), 팻 클라이언트 관리자(Fat client Manager), 분석 관리자(Analysis Manager), 저장소로 구성되어 있다.

팻 클라이언트는 상황인지에 필요한 비정형 데이터를 수집하는 데이터 수집 모듈(Data collecting)과 수집된 데이터를 분석 모델이 사용할 수 있는 입력의 형태로 전처리 작업을 수행하는 데이터 전처리 모듈(Data preprocessing)로 구성되어 있다. 팻 클라이언트는 인스턴스 형태이며 디바이스 계층의 디바이스와 수집되는 비정형 데이터의 종류에 따라 여러 개 존재한다.

파서는 추상화 과정의 수집, 전처리 과정에 필요한 정보를 포함하는 팻 클라이언트 프로파일과 분석 과정에 필요한 정보를 포함하는 AMDL 프로파일을 각 관리자가 사용할 수 있도록 파싱하여 분석 모델 구성 파일(Analysis Model Configuration File)과 팻 클라이언트 구성 파일(Fat client Configuration File)로 제공한다.

팻 클라이언트 관리자는 파서에서 제공하는 팻 클라이언트 구성 파일을 통해 팻 클라이언트를 생성하는 역할을 수행하는 팻 클라이언트 생성 모듈(Fat client Generating)로 구성되어 있다.

분석 관리자는 비정형 데이터의 상황 정보 추상화 과정의 분석과정을 수행하여 상황 정보로 추상화 하는 역할을 수행한다. 분석 관리자는 파서에서 제공하는 분석 모델 구성 파일을 통해 분석 모델의 실행환경을 구성한다. 그리고 분석 과정

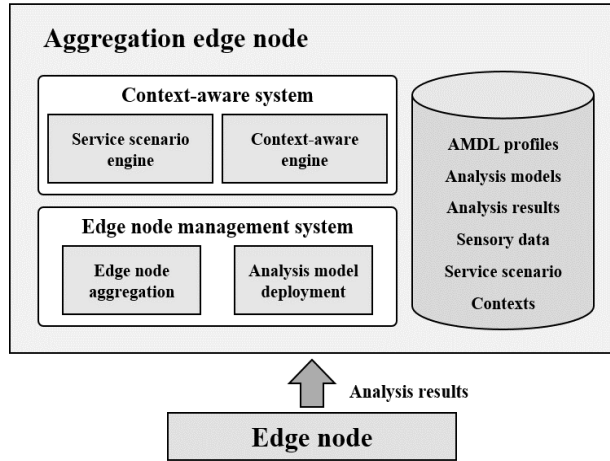


Fig. 3. System Architecture of Aggregation Edge Node for Context-Aware Services

을 수행하는 분석 모듈(Analysis)로 구성되어 있다. 저장소는 에지 노드 영역에서 수집되는 모든 센서 데이터와 팻 클라이언트 프로파일, AMDL 프로파일, 분석 모델을 포함한다.

2) 통합 에지 노드

통합 에지 노드는 에지 노드에서 비정형 데이터 추상화 과정의 분석을 위한 분석 모델의 배포 및 관리와 에지 노드에서 추상화된 상황 정보를 종합하여 상황인지 서비스를 제공한다. 통합 에지 노드의 자세한 구조는 Fig. 3과 같다. 구성요소는 에지 노드 관리 시스템(Edge Node Management System), 상황인지 시스템(Context-aware System), 저장소로 구성되어 있다.

에지 노드 관리 시스템은 비정형 데이터 추상화의 분석 과정에서 사용되는 학습된 딥러닝 모델을 에지 노드에서 사용할 수 있도록 배포 및 관리하는 분석 모델 배포 모듈(Analysis Model Deployment)과 에지 노드에서 추상화된 상황 정보를 종합하는 역할을 수행하는 에지 노드 통합 모듈(Edge Node Aggregation)로 구성되어 있다. 상황 정보 추상화의 핵심은 딥러닝 모델을 통한 환경 데이터 분석에 있다. 이러한 분석 모델의 정확도를

유지하기 위해 에지 노드 관리 시스템은 클라우드 계층에서 학습된 분석 모델을 주기적으로 업데이트 및 관리한다.

상황인지 시스템은 추상화된 상황 정보를 바탕으로 현재 상황을 판단하는 상황인지 엔진(Context-aware Engine)과 상황에 맞는 서비스 시나리오에 따라 적절한 서비스를 제공하는 서비스 시나리오 엔진(Service Scenario Engine)으로 구성되어 있다.

저장소는 에지 노드에 배포 및 관리하는 분석 모델과 AMDL 프로파일, 에지 노드 프로파일, 분석 결과, 수집 데이터, 서비스 시나리오, 상황 정보를 포함한다.

3.3 프로파일

비정형 데이터를 상황 정보로 추상화 할 때 수집-전처리-분석 과정을 지시하고 관리하기 위한 팻 클라이언트 프로파일과 AMDL 프로파일을 제공한다. 팻 클라이언트 프로파일과 AMDL 프로파일은 sensor data acquisition model schema[23]와 ML-Schema[24]를 기반으로 구성되었으며 상세한 내용은 다음과 같다.

1) 팻 클라이언트 프로파일

팻 클라이언트 프로파일의 개요는 Fig. 4와 같다. 수집 및 전처리 과정에 필요한 수집 데이터 정보 및 전처리 데이터 정보를 포함하고 있으며 구성 요소는 크게 5개의 상위 요소로 구성되어 있다.

hasInterface 요소는 데이터 수집 디바이스의 외부 인터페이스 사용 유무를 확인하는 요소이다. entityName 요소는 생성되는 팻 클라이언트의 고유 이름을 나타낸다. connectionType 요소는 센서 디바이스에 접근하기 위해 필요한 사용 프로토콜, IP주소 등 연결에 대한 명세를 나타낸다. gatherValue 요소는 센서 데이터를 수집할 때 사용되는 요소로 수집되는 데이터 타입, 파일 형식, 수집 데이터 차원 등 실제 수집된 센서 데이터의 명세를 나타낸다. preprocessedValue 요소는 수집된 데이터가 분석 모델의 입력으로 사용하기 위해 어떤 형식으로 처리되어야 하는지 전처리 후의 데이터 정보를 포함하

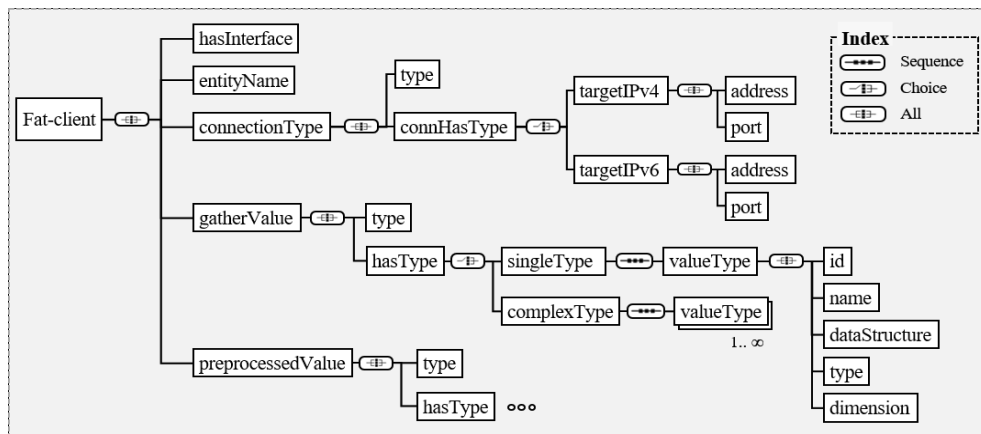


Fig. 4. Overview of Fat Client Profile for Providing Scalability to Fat Client

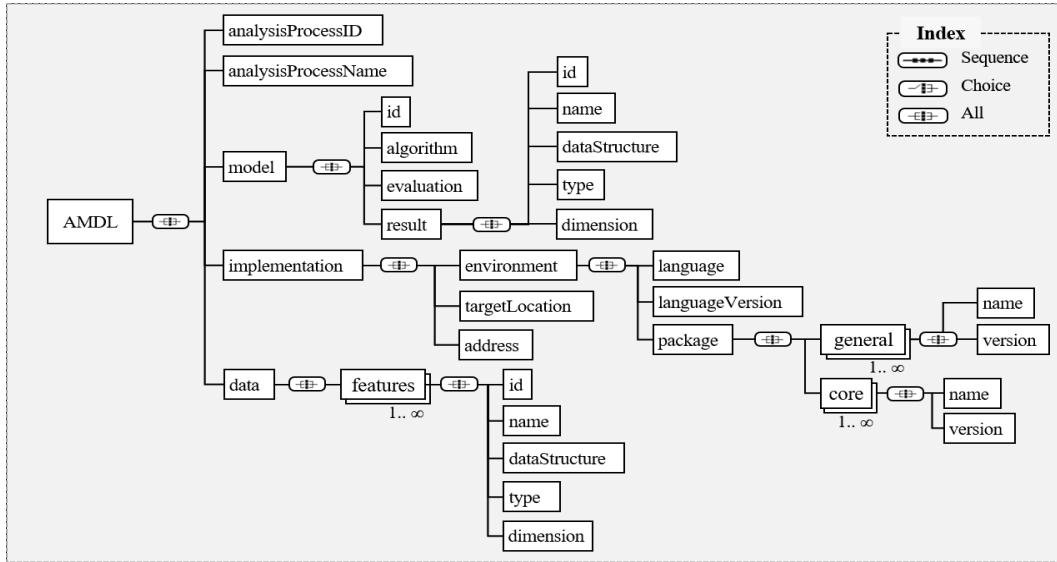


Fig. 5. Overview of AMDL Profile for Configuring Execution Environment of Analysis Model

고 있으며 이 정보들은 AMDL 프로파일의 분석 모델 입력 정보를 기반으로 작성된다. 팻 클라이언트 프로파일을 통해 비정형 데이터 추상화 과정의 수집과 전처리 과정을 관리하여 서비스 추가 및 변경에 따른 디바이스 추가 및 변경에 대한 시스템의 기능적 확장성을 제공한다.

2) AMDL(Analysis Model Description Language) 프로파일

AMDL 프로파일의 개요는 Fig. 5와 같다. 상황 정보 추상화의 분석 과정에 필요한 분석 모델의 실행 환경을 구축하기 위한 정보와 입출력 정보를 포함하고 있으며 구성요소는 크게 5개의 상위 요소로 구성되어 있다.

analysisProcessID와 analysisProcessName 요소는 분석 모델을 식별하기 위한 식별자이다. model 요소는 분석 모델의 정보를 기술하기 위한 요소를 갖는다. 이에 model 하위 요소는 사용되는 알고리즘을 표현하며 분석 모델의 출력 결과와 하위 요소인 evaluation 요소를 기반으로 분석 모델의 정확도를 평가한다. implementation 요소는 분석 모델의 사용 언어, 종속 패키지 정보, 분석 모델의 경로, 분석 결과 제공 목적지 등 분석 모델의 실행환경을 구축하기 위한 정보를 포함하고 있으며 기술된 정보를 기반으로 비정형 데이터 추상화의 분석 과정을 진행한다. data 요소는 분석 작업에 필요한 입력 데이터 정보를 포함하고 있으며 기술된 정보는 팻 클라이언트 프로파일의 전처리 데이터 정보에 사용된다.

3.4 팻 클라이언트 기반 비정형 데이터 추상화 방법

에지 컴퓨팅 환경에서 딥러닝 기술을 활용한 비정형 데이터 추상화 방법은 시스템의 기능적 확장성에 대한 요구사항에 따라 추상화 과정을 분리하여 수행하고 관리할 수 있어야 한다. 이에 본 논문은 팻 클라이언트 기술을 활용하여 추상화 과정의 수집과 전처리 과정을 인스턴스화하여 분석 과정과

분리하여 수행한다. 이러한 추상화 방법은 분리된 추상화 과정을 관리할 수 있어야 하기 때문에 팻 클라이언트 프로파일과 AMDL 프로파일을 제공하여 추상화 과정을 분리하여 관리한다. 비정형 데이터의 추상화 과정은 센서 디바이스에서 발생하는 환경정보를 팻 클라이언트에서 제공하는 팻 클라이언트 프로파일의 정보들을 기반으로 수집하고, 분석 모델이 사용할 수 있는 형태로 전처리 작업을 수행한다. 그 후, AMDL 프로파일의 정보들을 기반으로 실행 환경이 구축된 분석 모델은 전처리가 끝난 환경정보를 딥러닝 기반 분석 과정을 거쳐 상황 정보로 추상화하게 된다.

에지 컴퓨팅 환경에서 비정형 데이터를 상황 정보로 추상화 할 때 본 논문에서 제안하는 방법을 사용한다면 추상화 과정을 분리하여 수행하고 관리할 수 있기 때문에 수집 디바이스의 변경이나 분석 모델의 실행환경 변화에 대한 상황인지 시스템의 기능적 확장성을 제공할 수 있다.

4. 실험

제안하는 방법을 통해 비정형 데이터가 상황 정보로 추상화되는 과정을 보이기 위해 실험에서는 에지 컴퓨팅 환경에서 차량 출입 통제 알림 서비스를 위한 차량 이미지 인식 모델[25]을 대상으로 팻 클라이언트 프로파일과 AMDL 프로파일의 예제를 통해 시스템의 기능적 확장성을 확인하고, 비정형 데이터의 수집, 전처리, 분석에 대한 추상화 과정별 모니터링을 진행한다.

4.1 실험 환경

에지 컴퓨팅 환경을 구현하기 위해 오픈스택(Openstack) 플랫폼을 이용하였으며 통합 에지 노드 1개 에지 노드 1개 디바이스 노드 1개로 구성하였다. 오픈스택은 컴퓨팅, 네트

Table 1. Elements of Openstack based Test Environment

	CPU or VCPUs	RAM	DISK	OS
Openstack Server	i7-9800x	64GB	4TB	Ubuntu 16.04 LTS
Aggregation Edge Node	3 VCPU	8GB	40GB	
Edge Node	2 VCPU	4GB	20GB	
Device Node	1 VCPU	2GB	20GB	

위킹, 저장소 및 이미지 서비스의 핵심 클라우드 인프라를 구축할 수 있는 가장 거대한 오픈 소스 프로젝트이다. 오픈스택은 저장소, CPU, RAM과 같은 리소스를 추상화해 하이퍼바이저를 통해 분할한 다음 필요에 맞게 배포한다. 오픈스택이 구현된 서버와 오픈스택 플랫폼을 이용하여 구현한 노드별 구현 및 테스트 환경은 Table 1과 같다.

비정형 데이터를 상황 정보로 추상화 하여 상황인지 시스템에서 사용하는 모습을 보이기 위해 본 연구에서는 차량 출입 통제 알림 서비스를 위한 차량 번호판 이미지 인식 모델을 한정으로 실험을 보인다. 차량 출입 통제 알림 서비스 내용은 다음과 같다.

- 1) 디바이스 노드에서 차량 이미지 정보를 제공한다.
- 2) 에지 노드의 팻 클라이언트에서 디바이스 노드에서 제공된 이미지의 수집 및 전처리 과정을 수행한다.
- 3) 에지 노드의 차량 번호판 이미지 인식 모델을 통해 이미지를 분석한다.
- 4) 차량 이미지 분석 결과는 출입하는 차량이 허가된 차량인지 판별하여 모니터링 정보로 제공한다.

본 실험의 대상인 차량 번호판 이미지 인식 모델은 차량 이미지를 입력받아 차량 번호판 정보를 제공한다. 이러한 분석 모델은 사용자의 서비스 요구사항에 따라 모델의 교체 및 업데이트 그리고 수집 디바이스의 교체로 인한 수집, 전처리 과정에 대한 수정이 가능해야한다. 이에 본 논문에서 제공하는 팻 클라이언트 프로파일과 AMDL 프로파일의 내용을 통해 시스템을 수정하는 개발자의 개발 확장성을 보인다.

4.2 프로파일 기반 상황인지 시스템의 기능적 확장성

본 절에서는 확장성 측면에서의 기존 시스템과 제안하는 프로파일 기반 시스템과의 차별성을 보이기 위해 AMDL 프로파일과 팻 클라이언트 프로파일의 예를 들어 설명한다. 팻 클라이언트 프로파일은 비정형 데이터 추상화 과정의 데이터 수집 및 전처리 작업을 수행하는 팻 클라이언트의 구성 및 관리에 필요한 정보를 포함하고 있는 프로파일이며 AMDL 프로파일은 분석 과정에 필요한 정보를 포함하고 있는 프로파일이다. 이러한 프로파일의 개요에 맞게 작성된 문서가 적용된 시나리오는 Fig. 6과 같다.

개발자는 사용자 요구 사항에 따라 분석 모델이 교체될 경우 분석 모델의 실행환경, 전처리 과정, 수집 과정 등 추상화 과정의 환경 변화에 대한 추가적인 작업을 수행해야 한다.

Fig. 6은 관련 연구 절에서 언급했던 기존 시스템에서의 추가 작업 부분인 'A' 파트와 본 논문에서 제안하는 방법을 적용한 시스템에서의 추가 작업 부분인 'B' 파트로 구성되어 있다. 새로운 분석 모델의 코드와 학습 가중치를 담고 있는 파일을 전달받은 'A' 파트의 개발자는 제공받은 분석 모델의 코드를 확인하여 필수 패키지를 다운로드해야 하며 언어의 버전에 따른 실행 환경을 재구성해야 한다. 또한 분석 모델의

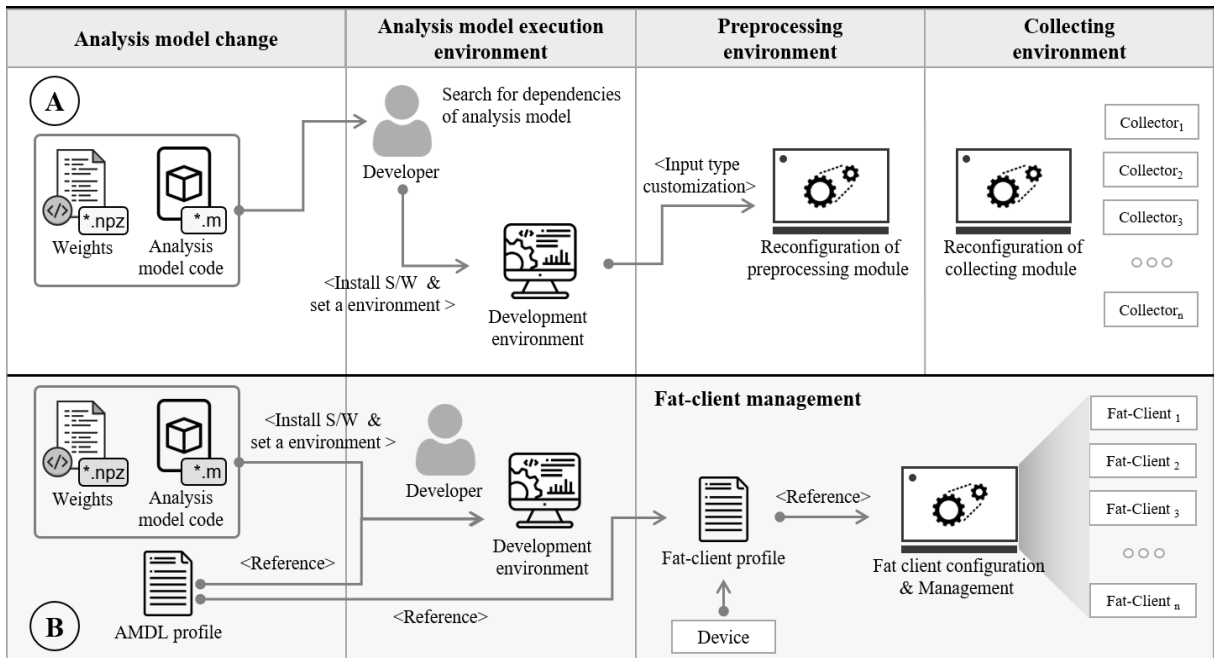


Fig. 6. Comparison of Abstracting Unstructured Data by Altering the Analysis Module of Existing and Proposed Approach

```

{
  "FatClient": [
    {
      "hasInterface": "false",
      "entityName": "Fat-Client Instance 1",
      "connectionType": {
        "type": "IPv4",
        "connHasType": {
          "targetIPv4": {
            "address": "172.24.4.14",
            "port": "8991"
          }
        }
      }
    },
    {
      "gatherValue": {
        "type": "singleType",
        "hasType": {
          "singleType": {
            "valueType": {
              "id": "sensor_data",
              "name": "car_plate_picture",
              "dataStructure": "jpg",
              "type": "image",
              "dimension": "658*468"
            }
          }
        }
      }
    },
    {
      "preprocessedValue": {
        "type": "singleType",
        "hasType": {
          "singleType": {
            "valueType": {
              "id": "input_1",
              "name": "car_plate_picture",
              "dataStructure": "jpg",
              "type": "image",
              "dimension": "650*460"
            }
          }
        }
      }
    }
  ]
  ... skipped ...
}
    
```

Fig. 7. Document Describing fat Client Instances based on Fat Client Profile

입력 형태에 맞춰 전처리 과정을 수행하는 전처리 모듈을 재구성해야 하며 수집 디바이스 교체에 대한 수집 모듈을 재구성해야 한다. 논문에서 제안하는 프로파일 기반 상황인지 시스템을 사용하는 'B' 파트의 개발자는 새로운 분석 모델의 코드와 학습 가중치를 담고 있는 파일과 함께 분석 모델의 원시 개발자가 작성한 AMDL 프로파일을 전달받는다. 'B' 파트의 개발자는 AMDL 프로파일을 참조하여 분석 모델의 실행 환경을 재구성할 수 있으며 Fat client 프로파일을 참조하여 수집, 전처리 과정을 수행하는 Fat client의 구성 및 관리가 가능하기 때문에 개발 확장성을 고려한 개발을 할 수 있다. 제

```

{
  "analysisProcessId": "analysis_proc03anpr",
  "analysisProcessName": "anpr_analysis01",
  "model": {
    "id": "stress_analysis_anpr02",
    "algorithm": "CNN",
    "evaluation": 0.88,
    "result": [ {
      "id": "output_1",
      "name": "car_plate_result",
      "dataStructure": "default",
      "type": "string",
      "dimension": "1*1"
    } ]
  },
  "implementation": {
    "environment": {
      "language": "python",
      "languageVersion": "3.7",
      "package": {
        "general": [ {
          "name": "opencv-python",
          "version": "4.2.0.34"
        },
        {
          "name": "numpy",
          "version": "1.18.1"
        }
      ],
      "core": [ {
        "name": "TensorFlow",
        "version": "1.14.0"
      } ]
    },
    "targetLocation": "remote",
    "address": "C:/document/workspace/anpr_v2.0.ml"
  },
  "data": [ {
    "id": "input_1",
    "name": "car_plate_picture",
    "dataStructure": "jpg",
    "type": "image",
    "dimension": "650*460"
  } ]
}
    
```

Fig. 8. Document Describing Number Plate Recognition Model based on AMDL profile

공하는 프로파일의 구체적인 예시는 Fig. 7, 8과 같다.

Fig. 7은 팻 클라이언트 프로파일 예제의 일부를 보여준다. 팻 클라이언트 프로파일은 entityName요소로 인스턴스 식별자를 가지고 있다. (A)영역은 수집 디바이스와의 연결 방식에 대한 정보를 포함하고 있으며 연결방식에 따라 하위 요소별 필요한 정보들이 기술되어있다. (B)영역은 수집 디바이스로부터 수집되는 데이터의 정보를 포함하고 있으며 수집

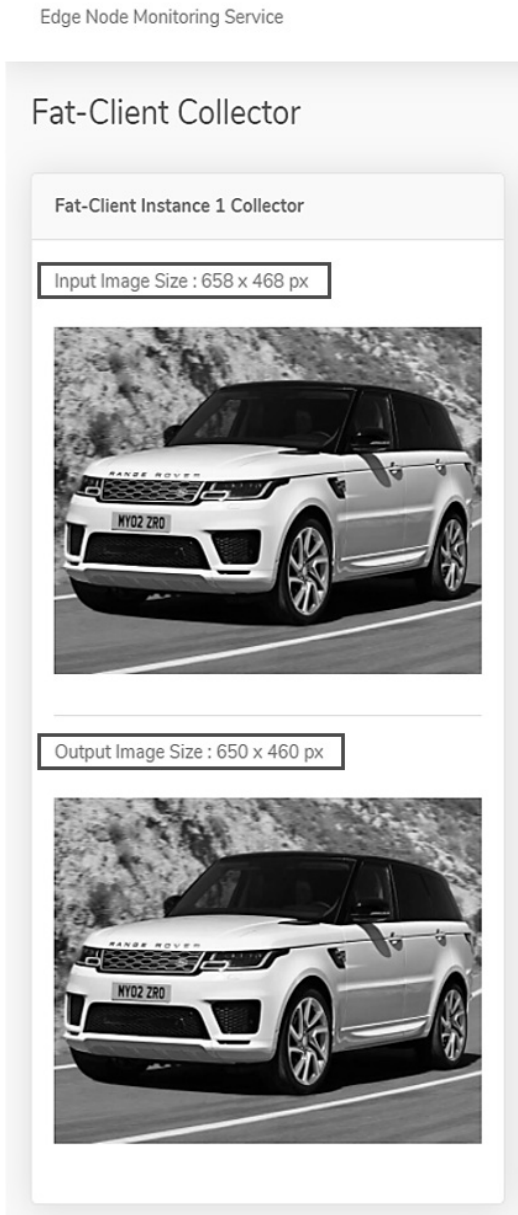


Fig. 9. Results According to the Collection and Preprocessing of the Fat Client Instance

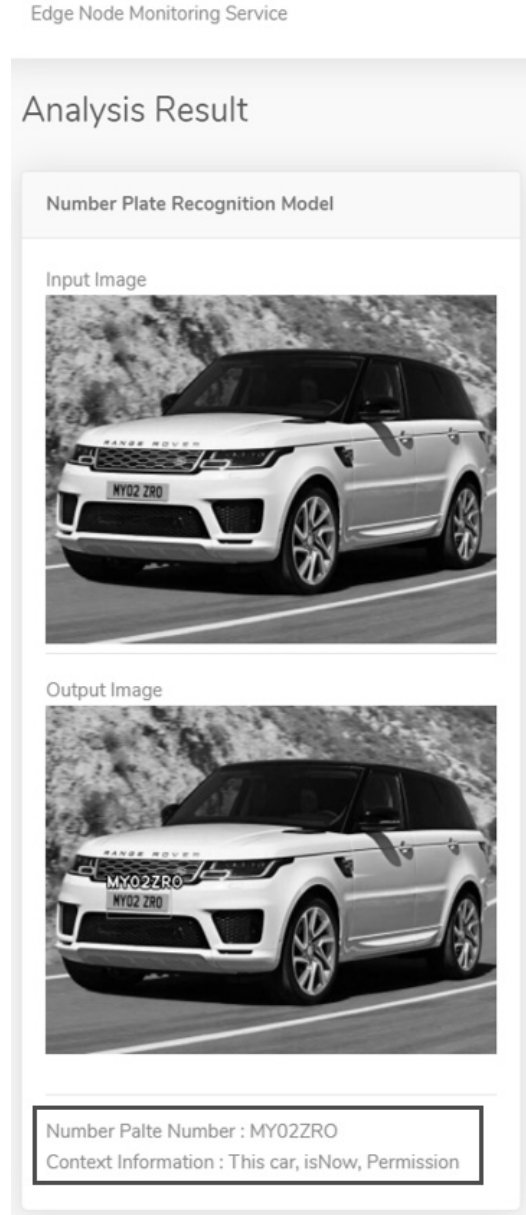


Fig. 10. Results According to the Input Image Abstraction based on Number Plate Recognition Model

되는 데이터 타입에 따라 필요한 자료구조 및 차원 등의 하위 요소가 기술되어 있다. 본 예제에 서 수집하는 데이터는 이미지 형태이며 'jpg' 파일 형식의 파일 포맷을 사용한다. (C)영역은 수집된 이미지 파일을 분석 모델의 입력으로 전처리하기 위한 정보를 포함한다. 시스템 개발자는 서비스 요구사항에 따라 수집 디바이스의 교체 및 추가를 위해 팻 클라이언트 인스턴스를 추가하고자 할 때, 팻 클라이언트 프로파일을 통해 시스템 개발을 지원받는다.

Fig. 8은 AMDL 프로파일의 예제의 일부를 보여준다. AMDL 프로파일은 최상위 요소로 분석 모델의 식별자를 가지진다. (A)영역은 본 예제에서 사용하는 차량 번호판 인식

모델의 알고리즘, 정확도, 실행 결과에 관한 정보를 나타낸다. 차량 이미지 인식 모델은 CNN알고리즘을 사용하며 최종적으로 문자열 형태의 출력 결과를 제공한다. (B)영역은 차량 이미지 인식 모델의 실행 환경 및 설정에 관한 정보를 나타낸다. 시스템 개발자는 (B)영역의 개발 언어, 언어의 버전, 패키지 등의 정보를 통해 필요한 소프트웨어를 설치하고 모델을 실행할 수 있는 환경을 구축한다. (C)영역은 분석 모델의 입력 정보를 나타낸다. 이 정보들은 분석 모델에서 사용할 수 있는 형태의 입력 정보를 포함하고 있으며 팻 클라이언트 프로파일의 전처리 정보에 사용된다. 시스템 개발자가 분석 모델의 추가 및 업데이트를 해야 할 때, AMDL 프로파일을 통

해 분석 모델의 실행 환경 구축을 지원받는다.

본 논문에서 제안하는 비정형 데이터 추상화 방법을 적용한 시스템은 추상화 과정이 프로파일을 통해 구성되고 관리되어 사용자의 요구 사항에 따라 다양한 도메인에서의 기능적 확장성을 제공하기 때문에 수집, 전처리, 분석의 비정형 데이터 추상화 과정의 종속성으로 인해 한정적인 도메인에서의 기능적 확장성을 제공하는 기존 시스템과의 차별성을 가진다. 이러한 차별성을 통해 시스템에서 사용하는 분석 모델의 교체, 디바이스의 교체에 따른 추가적인 작업이 발생하였을 때 제공하는 프로파일을 통해 시스템의 기능 추가 및 유지보수에 들어가는 개발자의 개발 확장성이 개선되었다.

4.3 비정형 데이터 추상화 과정

팻 클라이언트 프로파일과 AMDL 프로파일을 통해 수집, 전처리, 분석 과정을 관리하고 비정형 데이터가 추상화 되는 과정을 모니터링을 통해 설명한다. 비정형 데이터 추상화 과정의 수집과 전처리 과정은 팻 클라이언트 인스턴스를 통해 진행되고 분석 과정은 분석 모델에 의해 진행된다. Fig. 9은 팻 클라이언트 프로파일을 통해 생성된 팻 클라이언트 인스턴스의 수집 데이터와 전처리 데이터의 모니터링 결과이다. 센서 디바이스로부터 수집된 차량 이미지가 분석 모델의 입력 형태로 전처리된 것을 확인할 수 있다. 수집과 전처리 과정이 끝난 차량 이미지는 차량 번호판 이미지 인식 모델을 통해 상황 정보 추상화 과정이 진행된다. Fig. 10은 전처리된 이미지를 입력으로 하여, 차량 번호판 영역을 검출하고 번호를 인식한 결과와 추출된 차량 번호를 상황 정보로 추상화한 결과이다.

5. 결 론

에지 컴퓨팅 환경에서 비정형 데이터를 활용한 상황인지 시스템의 기능적 확장성을 제공하기 위해서는 수집, 전처리, 분석의 추상화 과정을 분리하여 수행하고 관리할 수 있어야 한다. 이에 본 논문에서는 수집, 전처리 과정과 분석 과정을 분리하여 수행할 수 있도록 팻 클라이언트 프로파일과 AMDL 프로파일을 제공하였다. 또한 에지 컴퓨팅 환경에서 팻 클라이언트 기반 비정형 데이터 추상화 과정을 제공하는 프로파일을 활용하여 수행하는 시스템을 제안하였다.

실험에서는 차량 출입 통제 서비스를 위한 차량 이미지 인식 모델을 대상으로 팻 클라이언트 프로파일과 AMDL 프로파일의 개요를 따르는 예제를 통해 시스템 개발자의 개발 과정의 확장성을 보였으며 이미지 데이터를 상황 정보로 추상화하는 과정을 확인하였다.

추가적으로 팻 클라이언트 프로파일의 정보를 자동으로 연결하여 팻 클라이언트 인스턴스를 생성할 수 있는 방법과 AMDL 프로파일을 통해 분석 모델의 실행환경을 자동으로 구축할 수 있는 방법을 고려하여 시스템을 개선한다면 보다 범용적인 사용이 기대된다.

References

- [1] R. Dobrescu, D. Merezeanu, and S. Mocanu, "Context-aware control and monitoring system with IoT and cloud support," *Computers and Electronics in Agriculture*, Vol.160, pp.91-99, 2019.
- [2] D. Preuveneers and E. Ilie-Zudor, "Big Data for context-aware applications and intelligent environments," *Future Generation Computer Systems*, Vol.99, pp.644-645, 2019.
- [3] L. C. Gubert, C. A. da Costa, and R. Rosa Righi, "Context awareness in healthcare: a systematic literature review," *Universal Access in the Information Society*, Vol.19, No.2, pp.245-259, 2020.
- [4] Z. A. Almusaylim and N. Zaman, "A review on smart home present state and challenges: Linked to context-awareness internet of things (IoT)," *Wireless Networks*, Vol.25, No.6, pp.3193-3204, 2019.
- [5] E. J. Kim, A. J. Jong, and N. S. Kim, "The method of providing IoE-based hierarchical context awareness," *2018 International Conference on Information and Communication Technology Convergence (ICTC)*, IEEE, 2018.
- [6] Kyungyong, Chung, Hyun Yoo, and Do-Eun Choe, "Ambient context-based modeling for health risk assessment using deep neural network," *Journal of Ambient Intelligence and Humanized Computing*, Vol.11, No.4, pp.1387-1395, 2020.
- [7] S. R. Pungus, J. Yahaya, A. Deraman, and N. H. B. Bakar, "A data modeling conceptual framework for ubiquitous computing based on context awareness," *International Journal of Electrical & Computer Engineering*, Vol.9, No.6, pp.5495-5501, 2019.
- [8] T. Hofer, W. Schwinger, M. Pichler, G. Leonhartsberger, J. Altmann, and W. Retschitzegger, "Context-awareness on mobile devices-the hydrogen approach," in *Proceedings of the 36th annual Hawaii International Conference on System Sciences*, 2003.
- [9] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge computing: Vision and challenges," *IEEE Internet of Things Journal*, Vol.3, No.5, pp.637-646, 2016.
- [10] Y. Wang, M. Liu, P. Zheng, H. Yang, and J. Zou, "A smart surface inspection system using faster R-CNN in cloud-edge computing environment," *Advanced Engineering Informatics*, Vol.43, No.101037, pp.1-9, 2020.
- [11] J. Ren, Y. Guo, D. Zhang, Q. Liu, and Y. Zhang, "Distributed and efficient object detection in edge computing: Challenges and solutions," *IEEE Network*, Vol.32, No.6, pp.137-143, 2018.

- [12] M. Nakatsugawa, et al., "The needs and benefits of continuous model updates on the accuracy of RT-induced toxicity prediction models within a learning health system," *International Journal of Radiation Oncology* Biology* Physics*, Vol.103, No.2, pp.460-467, 2019.
- [13] H. Miao, A. Li, L. S. Davis, and A. Deshpande, "Towards unified data and lifecycle management for deep learning," in *Proceedings of the IEEE 33rd International Conference on Data Engineering (ICDE)*, 2017.
- [14] H. J. Jeong, K. S. Park, and Y. G. Ha, "Image preprocessing for efficient training of yolo deep learning networks," in *Proceedings of the IEEE International Conference on Big Data and Smart Computing (BigComp)*, 2018.
- [15] S. K. Kim and J. H. Huh, "A study on the LMS platform performance and performance improvement of KMOOCSs platform from learner's perspective," *Journal of Ambient Intelligence and Humanized Computing*, pp.1-20, 2018.
- [16] S. Raza, S. Wang, M. Ahmed, and M. R. Anwar, "A survey on vehicular edge computing: Architecture, applications, technical issues, and future directions," *Wireless Communications and Mobile Computing*, Vol.19, No.3159762, pp.1-19, 2019.
- [17] L. Liu, C. Chen, Q. Pei, S. Maharjan, and Y. Zhang, "Vehicular edge computing and networking: A survey," *Mobile Networks and Applications*, pp.1-24, 2020.
- [18] H. El-Sayed and M. Chaqfeh, "Exploiting mobile edge computing for enhancing vehicular applications in smart cities," *Sensors*, Vol.19, No.5, pp.1073, 2019.
- [19] M. Z. Uddin, "A wearable sensor-based activity prediction system to facilitate edge computing in smart healthcare system," *Journal of Parallel and Distributed Computing*, Vol.123, pp.46-53, 2019.
- [20] M. Chen, W. Li, Y. Hao, Y. Qian, and I. Humar, "Edge cognitive computing based smart healthcare system," *Future Generation Computer Systems*, Vol.86, pp.403-411, 2018.
- [21] M. P. Hosseini, T. X. Tran, D. Pompili, K. Elisevich, and H. Soltanian-Zadeh, "Deep learning with edge computing for localization of epileptogenicity using multimodal rs-fMRI and EEG big data," in *Proceedings of the IEEE international conference on autonomic computing (ICAC)*, 2017.
- [22] S. Tuli, et al., "Healthfog: An ensemble deep learning based smart healthcare system for automatic diagnosis of heart diseases in integrated iot and fog computing environments," *Future Generation Computer Systems*, Vol.104, pp.187-200, 2020.
- [23] Y. S., Park, J. S., Choi, and J. Y. Choi, "Heterogeneous Sensor Data Acquisition Model for Providing Healthcare Services in IoT Environments," *KIPS Transactions on Software and Data Engineering*, Vol.6, No.2, pp.77-84, 2017.
- [24] G. C. Publio, et al., "ML-Schema: Exposing the Semantics of Machine Learning with Schemas and Ontologies," in *Proceedings of the 2nd Reproducibility in Machine Learning Workshop at ICML*, Stockholm, Sweden, 2018.
- [25] Matthew Earl, "Using neural networks to build an automatic number plate recognition system," GitHub, August 30, accessed Jun. 30, 2020, <https://github.com/matthewearl/deep-anpr>.



김도형

<https://orcid.org/0000-0002-8951-8679>
 e-mail : dohyungkim@soongsil.ac.kr
 2019년 한경대학교 컴퓨터공학과(학사)
 2019년 ~ 현 재 숭실대학교 컴퓨터학과 석사과정
 관심분야 : 클라우드 컴퓨팅, 빅데이터 처리, 인공지능 모델링



문종혁

<https://orcid.org/0000-0002-4919-0058>
 e-mail : jonghyeokmun@soongsil.ac.kr
 2017년 한경대학교 컴퓨터공학과(학사)
 2019년 숭실대학교 컴퓨터학과(석사)
 2019년 ~ 현 재 숭실대학교 컴퓨터학과 박사과정
 관심분야 : 로봇 미들웨어, 빅데이터 처리, MLOps, XAI



박유상

<https://orcid.org/0000-0002-7606-820X>
 e-mail : yspark@soongsil.ac.kr
 2014년 동아대학교 컴퓨터공학과(학사)
 2016년 숭실대학교 컴퓨터학부(석사)
 2016년 ~ 현 재 숭실대학교 컴퓨터학과 박사과정
 관심분야 : 시스템소프트웨어, 병렬/분산처리, 지능형 로봇



최 종 선

<https://orcid.org/0000-0001-9648-0667>
e-mail : jongsun.choi@ssu.ac.kr
2000년 송실대학교 컴퓨터학부(학사)
2002년 송실대학교 컴퓨터학부(석사)
2008년 ~ 2010년 유한대학교 e-비즈니스과
전임교원

2010년 송실대학교 컴퓨터학부(박사)
2011년 ~ 2012년 송실대학교 지능형로봇연구소 연구원
2012년 ~ 2013년 서일대학교 인터넷정보과 전임교원
2013년 ~ 현 재 송실대학교 컴퓨터학부 부교수
관심분야: 로봇 소프트웨어 플랫폼, 클라우드 컴퓨팅



최 재 영

<https://orcid.org/0000-0002-7321-9682>
e-mail : choi@ssu.ac.kr
1984년 서울대학교 제어계측공학과(학사)
1986년 미국 남가주대학교 전기공학과
(컴퓨터공학)(석사)
1991년 미국 코넬대학교 전기공학부
(컴퓨터공학)(박사)

1992년 ~ 1994년 미국 국립오크리지연구소 연구원
1994년 ~ 1995년 미국 테네시 주립대학교 연구교수
1995년 ~ 현 재 송실대학교 컴퓨터학부 교수
관심분야: 로봇 미들웨어, 시스템소프트웨어, 병렬/분산처리,
고성능컴퓨팅