

IoMT 표준 기반 Object Detection 서비스 제공을 위한 미디어 분석 서비스 운용 기술

*금승우 **문재원 ***김영기

전자부품연구원

*swkum@keti.re.kr

Design and Implementation of Object Detector based on IoMT Standard

*Kum, Seung Woo **Moon, Jaewon ***Kim, Youngkee

Korea Electronics Technology Institute

요약

최근의 IoT 기술의 발전은 다양한 형상, 네트워크 특징 및 서비스 아키텍처를 가지는 IoT 기기, 서비스 및 단말을 활용한 서비스의 발전을 가져오고 있다. 특히 OneM2M, OCF 등의 표준기구들은 다양한 IoT 기기 및 서비스 아키텍처에 대한 정의를 최근 수년간 진행해 오고 있으며, 이러한 IoT 서비스는 단순히 기기의 원격 상태 확인 및 제어 뿐만 아니라, 클라우드 및 AI 기술과의 연계를 통하여 그 서비스 영역을 지속적으로 확장 중에 있다. 이 중 Internet of Media Things 표준은 다양한 미디어 기반 서비스를 Thing으로 표현하여 다양한 Thing과의 연계 방안을 제시하고 있다. 본 논문에서는 기존에 다양한 기법을 통하여 연구 및 구현되고 있는 영상 기반 서비스를 Internet of Media Things 표준 기반으로 구현하기 위한 방법을 제시한다. 기존 영상 분석 기술은 대부분 정확도의 향상에 그 목적을 가지고 있어 서비스 형태로 제공하고 타 기기와의 연계성을 제공하기 위한 추가적인 기술간 연계가 필요하다. 본 논문에서는 Yolo v3 기반의 Face Detection 기술에 대하여, 해당 기술을 Internet of Media Things 표준으로 표출하기 위한 요구사항을 파악하고 실제 구현하기 위한 방법에 대하여 검토한다.

1. 서론

최근 수년간 이슈가 되고 있는 Internet of Things (IoT) 기술은 각 기기를 Thing으로 정의하고, 해당 Thing에서 제공할 수 있는 기술을 다양한 연결성 환경에서 정형화된 방법으로 표출하여 각 기기간 통신을 통한 서비스를 구성할 수 있는 기술로써, 현재 OneM2M [1] 및 OCF [2] 등의 표준을 통하여 호환성을 확보하고 서비스를 구성할 수 있도록 진행되고 있다. 최근의 IoT는 단순히 IoT 기술을 통한 Thing의 상태 확인 및 원격 제어, 단순 연계 수준의 서비스를 넘어서, Cloud 및 AI와 연계하여 보다 풍성한 서비스의 제공으로 그 영역이 확장되고 있다. 최근 출시되고 있는 AI Speaker들은 다양한 IoT 기기와의 연결성을 제공하고 있으며, 이러한 연결성을 기반으로 사용자의 음성명령을 분석, IoT 기기에 대한 제어를 제공한다. 또한 이와 같이 각 IoT 기기의 상태를 확인하고 분석하기 위해서는 단순 제어 명령 뿐 아니라 더 많은 정보를 포함하고 있는 오디오 및 비디오에 대한 분석이 반드시 요구되고 있으며, 이러한 오디오 비디오 데이터의 분석을 IoT 기술에 연계하기 위한 연구도 함께 진행되고 있다. 특히, ISO/IEC SC29 WG11에서는 이러한 미디어 분석 및 연계 기술의 IoT화를 위한 표준을 진행 중에 있으며, 해당 표준은 ISO/IEC 23093 Information technology - Internet of Media Things [3] 표준으로 제정이 진행 중에 있다. 본 논문에서는 단순 제어 뿐 아니라 AI, ML 등 보다 고성능의 서비스를 IoT 화하여 제공하기 위한 방법에 대하여 검토하고, 대중적으로 사용되고 있는 Yolo v3 [4] 기반의 Object Detection에 대한 구

현 방안을 IoMT 표준과 연계하여 제시한다.

2. Internet of Media Things 기반 Yolo v3

ML 기반 Object Detection은 최근 급격한 발전을 이루어 왔다. 그 중 Yolo 알고리즘은 영상에 대한 반복 처리 부분을 단순화하여 상대적으로 낮은 컴퓨팅 자원을 활용하여 높은 수준의 Object Detection을 제공하고 있다. Yolo v3는 Deep Learning을 활용하고 있으며, 이를 위하여 뉴럴 네트워크를 구성하고 학습 및 추론을 진행하도록 구성되었다. 하지만, 이러한 Yolo v3의 구성은 단일 머신 상에서 알고리즘의 정확도를 개선하기 위한 연구에 최적화되어 있어 인터넷 상에서 해당 서비스를 제공하기 위해서는 몇 가지 추가 기능이 함께 연동될 수 있어야 한다. 이는 단순히 Yolo v3의 문제는 아니며, 일반적으로 영상 기반 처리를 지원하는 딥러닝 알고리즘 개발환경들이 공통적으로 가지는 문제라고 볼 수 있다. 미디어 기반 딥러닝 분석 플랫폼들은 통상 로컬 머신에 저장되어 있는 영상을 뉴럴 네트워크에 적용하여 그 결과물을 도출한다. 이 과정에서 영상에 대한 선처리 과정이 통상적으로 요구되고 있다. 또한 최종적으로 도출되는 뉴럴 네트워크의 추론 결과물은 역시 로컬 머신에서 원본 이미지에 overlay되어 로컬 머신의 디스플레이에서 표현된다. 이러한 딥러닝 기반 영상 처리 기술 개발환경은 로컬 머신 외에 서비스를 제공해야 하는 IoMT 단말에 적용하기에는 부적절하다. 우선, 제공되는 기술 (이 경우는 object detection)이 외부 인

터페이스를 통하여 제공되지 못하고 로컬에서만 제공되고 있으며, 결과물의 확인도 로컬 머신에서만 확인이 가능하다. 하지만, 실제로 본 Object Detection 서비스가 IoMT화 될 경우 서비스의 구성은 로컬 머신이 아닌 외부와의 연계를 통하여 제공되어야 한다. 예를 들어 입력의 경우도 저장된 파일이 아닌 실제 IP Camera로부터 입력되는 실시간 스트리밍을 활용해야 하며, 또한 분석된 결과도 로컬 머신의 디스플레이가 아닌 원격 머신 혹은 컴퓨팅 자원의 디스플레이로 제공될 수 있어야 한다. 상기와 같은 문제를 해결하여 학습된 모델을 활용한 서비스가 IoMT에서 보다 다양하게 활용될 수 있게 하기 위하여, 본 논문에서는 Object Detection을 모듈러하게 구성하고 이를 IoMT 서비스로 표출하기 위한 단위 기술들을 함께 배치하는 방안을 제시한다. 그림 1은 본 논문에서 제시하고 있는 Yolo v3 기반 Object Detection의 IoMT MAnalyzer 서비스 모듈 구성도이다.



그림 1 Yolo v3 IoMT Analyzer 구성도

각 모듈의 정의는 다음과 같다.

Input Handler/ Preprocessor : 입력 미디어의 처리를 진행한다. 외부 미디어의 경우 미디어 핸들링의 작업과 함께, 입력된 미디어를 ML 모델에 적용하기 위한 선처리 과정을 진행한다.

Yolo v3 : Yolo 기반의 ML Computation을 진행한다. 선처리된 미디어를 Yolo Model의 input layer에 적용하여 연산을 수행하고 해당 모델의 Output Layer로부터 score를 취득한다.

Media Encoder : 입력 미디어에 대하여 yolo v3 결과를 overlay 한다. detection된 object에 대한 bounding box 및 클래스 명 표현이 이루어진다.

Media Streamer : ML 결과물에 대한 원격 확인을 위하여 결과물의 Streaming을 지원한다.

External Interface : 다른 MThing으로부터 상기의 과정에 대한 제어를 지원하기 위한 External Interface로 구현된다.

3. Implementation

상기와 같이 구성된 Yolo에 대하여 실제 구현을 통한 적용 가능성 여부가 검토되었다. Linux Box 기반으로 구현되었으며, 구현 환경은 Python3 기반의 Tensorflow 환경을 적용하였다. 각 기능은 별도의 python module로 구현되었으며, 외부 인터페이스의 연동을 위해서는 Flask 가 활용되었다. 각 기능들은 Flask를 기반으로 정의된 RESTful Interface를 통하여 외부 기기로부터 제어될 수 있다. 각 모듈의 구현을 위하여 Input handler/Preprocessor, Media Encoder는 OpenCV를 활용하였으며 Media Encoder는 OpenCV를 통하여 인코딩된 JPEG 이미지를 HTTP 기반으로 제공할 수 있는 Motion - JPEG으로 구성되었다. 그림 2는 상기와 같은 구현을 통하여 웹 브라우저로부터 확인되는 구현예이다. 외부 웹 브라우저는 MacOS의 Safari가 사용되었으며, 별도의 모듈이나 확장자의 설치 없이 MJPEG 분석 스트림을 실시간으

로 전송받아 표현하고 있다.

개발 환경에서는 별도의 GPU를 사용하지 않고 Intel i7 QuadCore 기반의 CPU를 적용하였으며, 1080i 입력 스트림에 대하여 약 4fps의 스트리밍이 가능함을 확인하였다.



그림 2 Yolo v3 구현 예

4. 결론

본 논문에서는 현재 개발 및 활용되고 있는 영상 기반 알고리즘을 IoMT 서비스로 제시하기 위하여 필요한 사항에 대하여 검토하고, 이에 대한 기본적인 기능의 구현이 진행되었다. 통상 로컬 머신을 기반으로 구성되어 있는 일반적인 개발환경에서의 문제점을 파악하고 이에 대하여 원격지로부터의 서비스 요청을 받아서 처리하고 제공하기 위한 서비스 구성안, 그리고 원격지로 분석 영상을 제공할 수 있는 방법에 대한 제시가 진행되었다.

본 논문에서 제시되는 방법은 Yolo v3 기반 object detection에 적용되고 있으나, 향후 타 영상/ 음성 기반 서비스에 손쉽게 확장이 가능하며, 이를 통한 마이크로 서비스화를 통한 서비스 구성 및 관리에 용이할 것으로 예상된다. 향후 클라우드 기술을 통한 현 모듈의 컨테이너화 및 마이크로서비스화, 그리고 정형적 인터페이스 정의 및 표준화에 대한 진행이 추가로 진행될 예정이다.

Acknowledgement

이 논문은 2019년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.1711082835, AI 어플리케이션을 지원하는 IoT 연동 분산 Edge-클라우드 기술 개발)

Reference

[1] OneM2M, <http://onem2m.org>.
 [2] Open Connectivity Foundation, <http://openconnectivity.org>.
 [3] ISO/IEC DIS 23093-2 Preview Information technology -- Internet of media things -- Part 2: IoMT Discovery and Communication API
 [4] J. Redmon and A. Farhadi. Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767, 2018.