**최종보고서**

**Machine Learning을 이용한**

**상황 판단 알림 시스템**

**(EDUM)**

**Ver. 1.3**

**2018.12.12**

**한국외국어대학교**

**정보통신공학과**

**3팀(EDUM)**

**문서 정보**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **구분** | **소속** | **성명** | **날짜** | **서명** |
| **작성자** | 한국외국어대학교 | 전진우 | 2018. 12. 12 |  |
| 한국외국어대학교 | 이대홍 | 2018. 12. 12 |  |
| 한국외국어대학교 | 임광효 | 2018. 12. 12 |  |
| 한국외국어대학교 | 권소연 | 2018. 12. 12 |  |
| 한국외국어대학교 | 김준영 | 2018. 12. 12 |  |
| **검토자** | 한국외국어대학교 | 전진우 | 2018. 12. 12 |  |
| 한국외국어대학교 | 이대홍 | 2018. 12. 12 |  |
| 한국외국어대학교 | 임광효 | 2018. 12. 12 |  |
| 한국외국어대학교 | 권소연 | 2018. 12. 12 |  |
| **사용자** | 한국외국어대학교 | 김준영 | 2018. 12. 12 |  |
|  |  |  |  |
|  |  |  |  |
| **승인자** | 한국외국어대학교 | 홍진표 |  |  |

머리말

본 문서는 Machine Learning을 사용하여 카메라의 영상에 대해 객체 인식을 한 후, 행동 인식을 통해 대단위 아파트, 상가단지 등과 같은 거주지역에서의 이상 상황 감지 및 알림을 전송하는 EDUM 시스템을 기술한 것이다.

**개정 이력**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **버전** | **작성자** | **개정일자** | **개정 내역** | **승인자** |
| 1.0 | 이대홍  전진우  김준영  임광효  권소연 | 2018. 12. 03 | 초안 작성 |  |
| **검토자** | 전진우 | | |
| 1.1 | 이대홍  전진우  김준영  임광효  권소연 | 2018. 11. 14 | 초안 수정 |  |
| **검토자** | 전진우 | | |
| 1.2 | 이대홍  전진우  김준영  임광효  권소연 | 2018. 12. 05 | 수정 보완 |  |
| **검토자** | 전진우 | | |
| 1.3 | 이대홍  전진우  김준영  임광효  권소연 | 2018. 12. 12 | 수정 보완 |  |
| **검토자** | 전진우 | | |

**목차**

**개요**

* 1. **목적**
  2. **관련문서**
  3. **용어 및 약어**

1. **제품소개**

**2.1 제품개요**

**2.2 제공 서비스**

1. **시스템 구성도**

**3.1 전체 시스템 구성**

**3.2 세부 시스템 구성**

**3.2.1 수집•제어부**

**3.2.2 객체•행동인식부**

**3.3 소프트웨어**

**3.4 하드웨어**

1. **시스템 상세설계**

**4.1 Filter**

**4.1.1 VideoCamera**

**4.1.2**

**4.2 Detection**

**4.2.1 Frame Receive**

**4.2.2 Object detection**

**4.2.3 Act Recognition**

**4.3 Server**

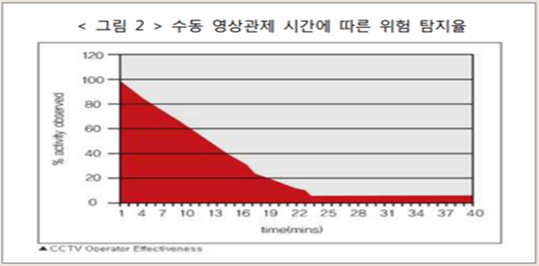
1. **제품기능 설명**
   1. **인터페이스 및 기능설명**
   2. **Class Diagram**
   3. **사용자 흐름도**
   4. **시나리오**
2. **실제적용방안 및 기대효과**
3. **프로젝트 세부추진계획 및 세부일정**
4. **Source Code**

**그림 목차**

**1. 개요**

본 장에서는 Inception v2 모델을 이용한 객체 인식 후, 행동 인식 알고리즘을 통하여 아파트 단지 내의 이상 상황을 서버로 전송하며, 서버에서는 관리자에게 SMS 및 알림 상황을 제공하는 웹사이트에 팝업 메시지를 발신하여, 신속한 대처가 가능하도록 하는 EDUM 시스템 및 구성, 기능에 대한 총괄 개요를 제공한다.

* 1. **목적**



[그림 321213213213]

기존의 보안 시스템은 관리자가 직접 영상을 보며 이상 상황을 판단하는 방식이다. 하지만 [그림]에 따르면, 지속적으로 사람이 직접 영상을 감시할 때, 시간이 지남에 따라 급격하게 관제 능력이 떨어짐을 알 수 있다. 따라서 본 프로젝트에서는 객체 인식을 통해 이상 상황을 판단하여 기존의 방식보다 더 효율적이고 기존의 방식보다 더 적은 비용으로 보안 시스템을 구축하는데 목적이 있다.

* 1. **관련문서**

|  |  |
| --- | --- |
| **문서** | **문서 제목** |
| 연구성과 실용화 진흥원 | 영상 감시 시스템 시장 및 기술동향 |
| 한국지역정보개발원 | 지능형 CCTV 기술 현황 및 활용 사례 |
| 한국디지털CCTV연구조합 | 차세대 지능형 CCTV 산업 경쟁력 강화 방안 연구 |
| 전자부품 연구원 | 지능형 CCTV 시스템 기술 이슈 및 산업동향 |
| Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun | Faster R-CNN: Towards Real-TimeObject Detection with Region Proposal Networks |
| Ross Girshick Jeff Donahue Trevor Darrell Jitendra Malik UC Berkeley | Rich feature hierarchies for accurate object detection and semantic segmentation |

**[Table 1] 관련문서**

* 1. **용어 및 약어**

|  |  |
| --- | --- |
| 용어 및 약어 | 풀이 |
| COCO dataset | Common object in Context dataset |
| SMS | Short Message Service |
| HTML | Hyper Text Markup Language |

**[Table 2] 용어 및 약어**

* Faster RCNN: Fast R-CNN의 Region Proposal 방법인 Selective Search 방식을 개선한 방식으로CNN을 통해 추출된 특징 맵을 RPN에 입력한다. RPN(Region Proposal Networks)에 입력 시 Object가 있을 만한 구역에 대한 Proposal을 연산한다.
* 가상 펜스: 접근 제한 구역에 가상으로 그어진 선으로 이 선을 넘어가는 사람을 감지하면 관리자에게 알림을 보낸다.

**2. 제품소개**

* 1. **제품개요**

EDUM은 관리 구역 내에서 발생하는 여러 이상 상황을 객체 인식과 상황 감지 시스템을 통해 알림을 제공하여 다수 카메라에 대한 편리한 관리 및 감독을 제공하는 시스템이다. 관리자는 이 알림으로 해당 구역의 실시간 영상을 확인하여 신속하게 현장에 출동할 수 있다.

* 1. **제공서비스**

관리자는 EDUM을 통해 여러 카메라를 한 화면에서 관제할 수 있으며, 그 영상들의 객체 인식 및 상황 감지 중 이상 상황이 감지 되었을 경우 그에 대한 알림을 웹 페이지의 팝업창과 SMS 알림을 통해 제공받을 수 있다. 이 알림에는 감지된 상황의 종 류 및 카메라에 대한 정보와 함께 해당 구역의 실시간 스트리밍 영상이 포함되어 있 다. EDUM은 주민에게도 제한된 구역에 한해 카메라 영상의 스트리밍 서비스를 제공하여 주민은 근처 장소의 카메라 영상을 볼 수 있다.

**3. 시스템 구성도**

* 1. **전체 시스템 구성**
  2. **세부 시스템 구성**
     1. **수집∙제어부(Camera Module, Filter)**

수집∙제어부는 사전에 고정한 FPS로 영상을 제공하는 각 구역에 설치된 카메라를 통하여 각 이미지 프레임을 객체∙행동인식부로 전송하는 기능을 수행한다.

Webcam을 통해 받은 영상은 Python, OpenCV 모듈을 사용하여 수집 및 처리하며, 각 webcam별로 수집된 영상의 프레임은 객체∙행동인식 결과에 따라 동적으로 바뀌는 스케줄에 따라 순차적으로 객체∙행동인식부로 전송한다.

* + 1. **객체∙행동인식부(Detection Module)**

객체∙행동인식부는 수신한 프레임을 통하여 해당 구역의 객체를 인식하고 객체의 상황을 파악하는 기능을 수행한다. Python, Tensorflow 모듈을 사용하여 프레임을 처리한다. 먼저 수집∙제어부로부터 수신 받은 이미지 프레임을 토대로 객체 인식을(사람, 쓰러진 사람, 쓰레기) 하여 결과를 이미지 프레임에 씌워 가공하며, 이후 가공한 이미지 프레임을 가지고 행동 인식(사람이 쓰러짐, 쓰레기 투기, 월담, 접근 제한구역 침입)을 한 후, 결과를 바탕으로 가중치를 계산하여 수집∙제어부에 전송한다. 이 계산된 가중치는 수집∙제어부가 보낼 카메라별 프레임 순서를 배치하는 스케줄을 구성할 때 사용된다. 또한 행동인식 결과를 Server로 전송한다. 객체를 인식하기 위해서는 객체 인식 모델이 필요하며, 사전 훈련된 모델이 제공하는 객체 이외의 객체를 인식하기 위해 이미지 샘플을 추가하여 머신 러닝을 수행한다. 추가 학습된 Custom Model을 기반으로 객체 인식을 수행하며, 객체 인식후에 각 객체의 상태와 좌표 값을 이전 프레임과 비교하여 사람이 쓰러졌는지, 쓰레기가 버려졌는지, 사람이 담을 넘는지, 제한 구역에 침입했는지를 확인한다. 그 후 확인된 결과를 토대로 데이터를 Server로 전송하게 된다.

* + 1. **Server & DataBase**

Server는 객체∙행동인식부로부터 전송받은 객체∙행동인식 결과 정보를 DB에 저장하며, 받은 정보를 바탕으로 Event를 발생시켜 handle 한다. Server는 객체∙행동인식 결과로부터 관리자가 확인해야 할 상황을 판단하여 DB에 카메라의 상태를 Warning 상태로 업데이트한다. 그리고 Warning 상태인 카메라가 감지될 시, Server는 관리자가 확인을 위해 띄워 놓은 브라우저에 메시지창을 팝업함으로써, 관리자에게 감지된 상황 및 감지된 곳의 위치를 알려주며, 감지된 장소 카메라의 영상을 스트리밍을 통하여 보여준다. 또한, 관리자의 부재를 대비하여, 현재 관리자 페이지에 로그인한 관리자의 정보 중 휴대폰 번호를 바탕으로 근무중인 관리자의 휴대폰에 SMS를 전송하여, 감지된 상황, 장소의 정보를 제공한다.

* 1. **사용한 소프트웨어**
* **Tensorflow**

Tensorflow는 구글이 개발하여 오픈소스로 공개한 기계학습 라이브러리이다. Window나 Linux등 다양한 OS에서 사용이 가능하다. 본 프로젝트는 위 라이브러리 중 Object Detection 파트를 중점적으로 사용하며, 카메라로 촬영한 영상을 처리하여 특정 객체를 인식하는 역할을 한다.

* **DataBase**

Django에서 제공하는 Sqlite3를 이용하여 서버와 연결하여 사용한다.

* **Server**

Django를 이용하여 웹서버를 제작하고 Daphne를 사용하여 웹서버를 배포한다.

* **UI**

웹 인터페이스는 HTML을 이용하여 페이지를 작성한다.

* **SMS**

CoolSMS에서 제공하는 API를 통하여 SMS을 사용자에게 전송한다.

* 1. **사용한 하드웨어**
     1. **웹캠 (Webcam)**

**아파트 단지 내에 설치하여 해당 구역의 상황을 촬영하는 카메라로, 영상을 끊김 없이 스트리밍해야 하기 때문에, 고성능 카메라인 C922 Pro Stream Webcam을 사용한다. 보조 카메라로** **SPC -A1200MB를 사용한다.**

**** 

**[Figure]**

**4. 시스템 상세설계**

1. **Filter**
   1. **VideoCamera**

컴퓨터에 연결된 camera를 실행시켜준다. 현재시간을 기준으로 output영상 객체를 카메라마다 각각 생성해주고 생성된 객체에 지속적으로 frame을 써주면 동영상으로 저장된다. 저장을 시작한 시간을 기준으로 사용자가 설정한 시간이 지나면 동영상 저장 객체를 중지시켜 저장을 완료한다. 그와 동시에 동영상 저장이 끊이지 않도록 즉시 output영상 객체를 재생성 후 새로 저장을 시작한다. 또한 모든 동영상 저장은 각 camera에 할당된 id를 통해 각각 저장된다.

* 1. **Frame sender**

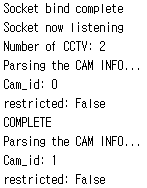
**VideoCamera**를 통해서 얻은 프레임을 객체∙행동인식부로 전송하는 클래스이다.

1. Detection

**[Figure] Detection Module 구성도**

* 1. **Frame Receive**

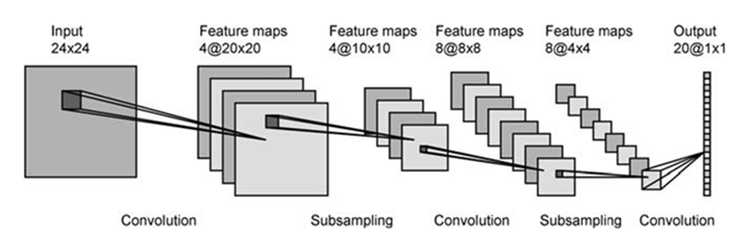
**Filter Module과 TCP 통신을 사용하여 데이터를 송수신하는데, Frame을 수신하기 전, Filter Module으로부터 연결된 카메라의 총 대수만큼 Cam 객체를 생성하고, 생성한 각 Cam 객체의 id 및 data를 조작하기 위해 리스트에 shallow copy하여 저장한다. 이후 Cam 객체의 id로 수신한 Frame을 parsing하며, Cam 객체의 frame 변수에 Frame을 저장 후 Object detection Module로 전달한다.**



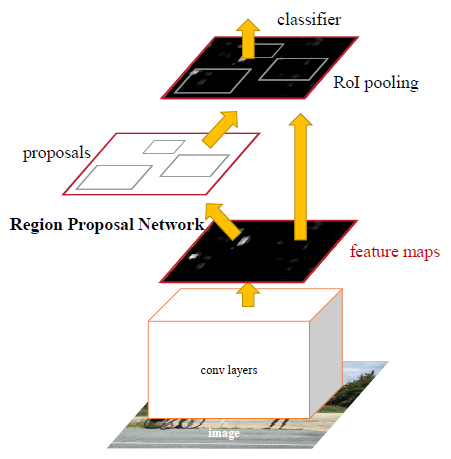
**[그림12312312] 2대의 카메라를 parsing하여 각 객체에 정보 저장**

**위와 같이 각 Cam 객체에 정보가 저장되어진다.**

* 1. **Object detection**

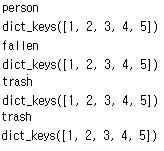
****

**[그림12312] CNN 연산 과정**

****

**[그림1231] Faster R-CNN 객체 추출 방식**

**객체 인식에 사용된 Inception V2 모델은 Faster R-CNN 알고리즘을 사용하여 객체를 인식한다. 비교 모델인 YOLO V3 모델이 더 빠른 연산을 제공하지만 Inception V2모델이 더 세밀한 객체의 인식률이 높다고 판단하여 사용하였다.**

****

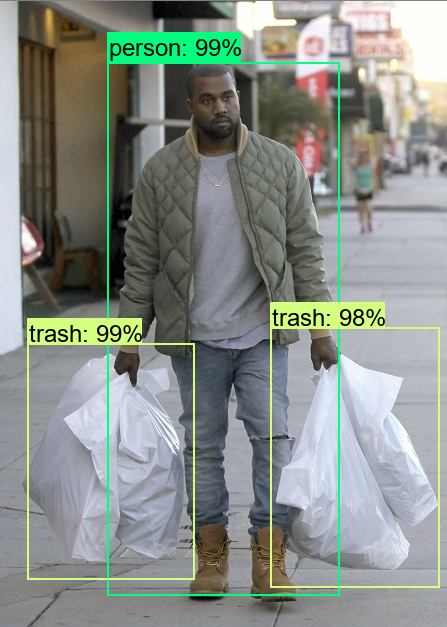
**[Figure] 객체 인식 결과**



**[Figure] 객체 인식 결과 저장**

**Frame Receive Module로부터 전달받은 Frame으로 객체인식을 수행한다. 객체는 person, fallen, trash, bottle, metal 5개가 있으며, 인식한 객체의 경계선, 정확도, class name과 인식한 객체의 개수가 변수에 저장된다.**

**이후 저장된 변수를 토대로 프레임 이미지에 씌워 가공하며, 가공 처리된 프레임을 Act Recognition Module로 전달한다.**

** **

**[Figure] 이미지 가공 처리**

* 1. **Act Recognition**

1. **Server**

**5. 제품기능 설명**