

## 스마트 관제를 위한 딥러닝 기반 이상행동 기술 동향 분석

\*이지애 \*\*문성철\*

서울기술연구원

\*jiaelee@sit.re.kr, \*\*sungchul.mun@sit.re.kr

### Brief Overview of Deep Learning based Anomaly Detection for Smart Surveillance System

\*Lee, Jiae \*\*Mun, Sungchul

Department of Smart City Research, Seoul Institute of Technology

#### 요약

스마트관제 시스템은 딥러닝 서버내 학습된 백본 네트워크 모델이 실시간으로 스트리밍 되는 CCTV 영상으로부터 이상행동 패턴을 선별적으로 탐지하고 관제요원에게 전달하여, 사전에 사건사고를 예방하거나 즉시 대응 체계의 유연한 운영을 가능케 하는 시스템이다. 최근 지능형 CCTV(Closed Circuit Television) 서비스가 일부 지역에 선별 관제의 형태로 시범적으로 운영되고 있는 상황이다. 지능형 시범서비스는 공공 영역에서 선별 CCTV 관제의 형태로 이상행동 상황을 즉각 인지하여 사건사고를 예방하거나 피해를 최소화하고자 하는 목적으로 주로 사용되고 있다. 그러나, 범죄 등의 특정 시나리오에만 한정해서도 이상 행동 유형이 너무나 다양하기 때문에 이상행동 영상의 사전분류(Annotation)를 통해 딥러닝 모델을 학습시키는 것이 현실적으로 어려운 상황이다. 따라서 본고에서는 최신 이상 행동 탐지(Anomaly detection) 알고리즘과 응용사례를 분석하여 실제 현장에 적용할 수 있는 현장 중심의 기법을 제안하고자 한다.

#### 1. 서론

스마트 관제 시스템(Smart Event Analysis System, SEAS)은 이상 행동패턴 실시간 감지를 통해 사전에 사건사고를 예방하거나 피해 정도를 최소화하고자 함을 목적으로 하는 스마트 시스템으로 정의된다. 이러한 스마트 관제 시스템은 국경보안, 해안, 도시가스 안전 등 일부 분야에서 현재 시범적으로 운영되고 있다. 그림 1에서 보는 바와 같이, 스마트 관제 시스템은 백엔드 서버, 네트워크, DBMS(Data management system) 및 분석부로 구분될 수 있다. 이러한 스마트 관제 시스템은 특정 이벤트를 시스템

분석부에서 선별적으로 탐지한 후 관제센터에 실시간 경고 알람 등을 제공하여 사건, 사고에 신속하게 대처할 수 있도록 하는 것을 목적으로 한다.

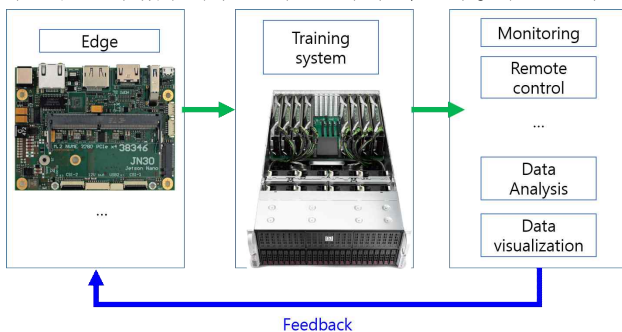


그림 1. 스마트 관제 시스템의 예시구조

Fig. 1. Example of Infrastructure in Smart Surveillance System

현재 CCTV(Closed-circuit television) 관제는 공공 영역에서 대기 환경 인식, 유량 인식, 범죄 상황 인식 등 다양한 분야에서 활용되고 있으나, 급증한 CCTV 수로 인해 1인 관제 요원이 모니터링 하는 대수는 지역에 따라 100대에서 많게는 600여대에 이르는 실정이다. 서울시 공공데이터 포털에 따르면 서울시 25개 자치구가 운영하는 CCTV 수는 약 5만 7천여대에 이르는 것으로 파악된다 [1]. 지능형 CCTV를 활용한 이상행동 탐지가 일부 지역에서 중요한 역할을 하고 있으나, 범죄 행동 등을 탐지할 수 있는 학습데이터가 충분치 않고, 기존 시스템에 비해 많은 하드웨어 리소스를 사용하기 때문에 다수의 영상 채널을 한꺼번에 분석하는 경우에 GPU 발열 등의 문제가 발생하고 있다. 따라서 적절한 범위에서 시스템 효율성을 보장할 수 있는 선별관제 시스템을 구축하는 것이 중요하다.

따라서 본고에서는 최신 이상 행동 탐지(Anomaly detection) 알고리즘과 응용사례를 분석하여 실제 현장에 적용 가능한 부분을 고찰하고, 이를 통해 향후 지능형 CCTV의 발전방향에 대해 논의하고자 한다.

#### 2. 이상행동 패턴 감지를 위한 기존 선행연구 분석

이상행동 패턴 감지를 위한 딥러닝 기반 선별관제 기술은 현재

YOLOv3 알고리즘과 Optical Flow 기반의 오브젝트 추적 기술이 주로 융합되는 추세이다. Optical Flow 기반의 방식은 특정 영상 프레임 t에서  $\Delta t$  만큼 변화했을 때의 픽셀단위의 Velocity 벡터 변화를 분석하여 이미지 프레임의 모션변화를 추정하는 방식을 주로 사용한다. 최근에는 이미지 데이터셋외에 이상행동 동영상 데이터 셋을 학습시켜 보다 다양한 상황의 이상행동 패턴을 감지하기 위한 딥러닝 연구들이 시작되는 단계로 볼 수 있다. Xiong [2]은 심층 메트릭 학습기반의 분류 알고리즘을 제안하여 다양한 환경에서의 동일 인물의 재식별화(Re-identification)를 가능케 하였다. Sharif [3]는 멀티 클래스 상관성과 유클리디안 유사도 기법을 활용하여 영상의 최적화된 특징을 추출하고 다중 클래스 SVM 분류 알고리즘으로 이상행동 패턴을 분류하는 방법을 제안하였다. Chandran [4]은 비재귀 모션 클러스터링 알고리즘 (Non-recursive motion similarity clustering algorithm)을 활용하여 최대 30명의 보행자를 인식하고 트래킹 할 수 있는 실시간 선별관제 시스템을 제안하였다. 최근 딥러닝 기반의 다양한 오브젝트 트래킹 알고리즘과 이상행동 패턴인식을 위한 다양한 딥러닝 백본 아키텍처[5]가 제안되고 있으나 실제 이상행동 패턴을 구분하기 위한 학습데이터 셋이 충분하지 않아 실제 현장에서 활용될 수 있는 지능형 선별관제 서비스는 특정 이벤트 상황에 한정하여 제안되고 있는 실정이다.

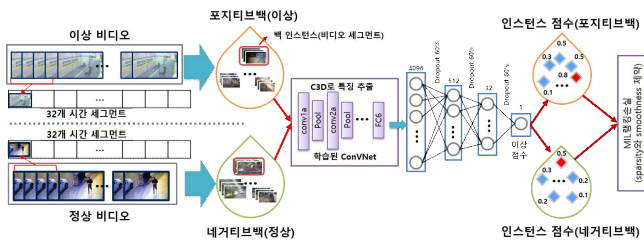


그림 2 [5]의 백본(back bone) 네트워크 구조

### 3. 이상행동 감지를 기존 오픈 데이터셋 분석

최근 급증한 CCTV 수에 대한 선별관제의 필요성이 대두되고 스마트시티 트랜스포메이션 추세에 따라 전세계적으로 오픈데이터 셋 구축이 시작되고 있는 단계이다. 다양한 오픈데이터 셋이 현재 공개되어 있는 상황이며 이용 가능한 이상행동 탐지를 위한 오픈 데이터 DB는 UMN, UCSD, Avenue, BOSS, Crowd, KTH, Weizmann, HMDB 51, UCF 데이터셋 등을 포함한다. UMN(University of Minnesota) 데이터셋은 5개의 스테이지로 구성되어 사람들이 걸어 다니거나 일정 시간이 지나면 다른 방향으로 달리는 비디오셋 등을 포함하며, 특정상황에서의 급격한 이동 등을 이상 행동으로 간주한다[4]. Li. 등은 UCSD(University of California San Diego)을 제안하였으며, Ped1과 Ped2 데이터셋을 나뉘며, 각각 70개와 28개의 감시 비디오 셋을 포함한다 [6]. 그러나 비현실적이고 단순한 행동을 이상 행동으로 간주하는 단점이 있다. Avenue 데이터셋은 37개의 비디오 셋을 포함하고 있으며, 많은 이상 행동들이 포함되어 있지만 비디오 클립이 짧고 한 곳에서만 영상이 녹화되어 있어 다양한 상황을 반영하지 못하는 한계점이 있다. BOSS 데이터셋은 기차에서 배우를 촬영한 감시 영상이며, 피로하거나 질병에 걸린 사람, 공항 상황 등을 이상행동으로 탐지하는 것을 특징으로 한다 [7]. Crowd 데이터 셋은 31개의 붐비는 관중을 포함한다 [8]하고, KTH 데이터셋은 박수, 걷기, 복싱, 달리기, 조깅 및 손 흔들기를 포함하여 6개 클래스로 구성되는 것을 특징으로 한다.

KTH 데이터셋은 4가지 시나리오에 대해 25명의 배우가 재현한 600개의 비디오 시퀀스를 포함한다. 2005년에 공개된 Weizmann 데이터셋은 달리기, 걷기, 흔들기 등을 포함한 10가지 행동 클래스로 90개의 비디오 시퀀스로 구성되었고, 9명의 배우에 의해 재현되었다. HMDB 51 데이터 셋은 푸시 업, 웨이브 등의 51가지 행동을 포함하며, 자전거 타기, 차기, 편지, 밀기, 던지기, 돌기 등을 포함한 12개 행동 범주를 포함하는 것을 특징으로 한다. 해당 데이터 셋들은 비디오 영상클립의 길이가 상대적으로 짧고 테스트 케이스가 적어 행동 변화를 추정하기에 많은 한계점을 지닌다. 또한 배우에 의해 인위적으로 재현되어 실제 상황을 적절히 반영하지 못한다는 한계점이 존재한다. Sultani 등 [9]은 CVPR 2018에서 체포, 방화, 강도, 폭행 등 총 13개의 카테고리 대해 이상행동에 반대되는 정상상황비디오 950개, 비정상 행동 패턴을 포함하는 950개 비디오를 수집하여 128시간 분량(95.9 GB)의 UCF (University of Central Florida) 오픈 데이터셋과 학습 네트워크를 공개하였다. UCF 데이터 셋은 현재까지 공개된 데이터 셋 중 가장 현실에서 일어날 특정 범죄 행동을 잘 반영할 수 있고, 이상행동 감지를 위해 학습 셋에 일일이 주석을 다는 과정을 최소화시킬 수 있다는 점에서, 공공 안전 분야에 시범적으로 활용될 수 있다고 판단된다.

### 4. 논의

본 논문은 스마트 선별 관제를 위한 최신 이상행동 패턴 탐지 알고리즘과 현재 활용 가능한 오픈데이터의 특징을 분석하여 현장에 시범적으로 적용 가능한 부분에 대해 고찰하고자 하였다. 최근 이상행동 감지를 위해 다중 인스턴스 학습에 기반한 이상행동 감지 기술 등이 제안되었지만, 다중채널의 CCTV를 딥러닝 기반으로 운영할 경우 GPU 발열 등의 문제로 인해 별도의 쿨링 시스템을 구축해야 하는 이슈가 발생할 수 있다. 또한 이상행동 패턴을 포함하는 동영상 자체를 학습시키는 최신 딥러닝 기술을 활용해 수백만 개의 이상행동 패턴을 학습한다 하더라도 실제로 발생할 수 있는 수많은 경우의 이벤트를 모두 포함하는 것은 불가능하기 때문에 특정 이벤트(방화, 폭력, 배회, 침입 등)에 한하여 선별관제 시스템을 시범적으로 운영하여 오탐율을 최소화하는 방향으로 우선 실증이 진행되어야 한다. 현재 이상행동 패턴 분류를 위해 활용 가능한 오픈 데이터셋도 제한적인 상황이라 KISA에서 일부 제공되고 있는 데이터셋외에도 활용 가능한 오픈 데이터 셋에 대한 통합표준화 작업이 필요하다. 가장 중요한 것은 기학습된 딥러닝 모델이 새로운 데이터셋이 구축되었을 때에도 기존 데이터베이스와의 유사성에 따라 전이학습이 가능하도록 설계되는 연구들이 많이 진행되어야 한다는 점이다. 또한 현장에서 효율적으로 운영될 수 있는 딥러닝 서버 스펙에 대한 표준가이드라인 도출과 음원분석 및 열감지 센서, IoT 센서들을 추가적으로 활용하여 오탐율을 최소화 할 수 있는 시스템 구성 및 실증 가능 서비스 방안에 대한 실효적인 가이드라인 도출에 대한 실증 연구가 우선적으로 수행되어야 한다.

### 참 고 문 헌 (References)

[1] 공공데이터 포털, <https://www.data.go.kr/>  
 [2] M. F. Xiong, D. Chen, J. Chen, J. Y. Chen, B. Y. Shi, C. Liang, and R. M. Hu, "Person re-identification with multiple

similarity probabilities using deep metric learning for efficient smart security applications," *Journal of Parallel and Distributed Computing*, Vol.132, pp.230-241, Oct 2019.

[3] A. Sharif, M. A. Khan, K. Javed, H. G. Umer, T. Iqbal, T. Saba, H. Ali, and W. Nisar, "Intelligent Human Action Recognition: A Framework of Optimal Features Selection based on Euclidean Distance and Strong Correlation," *Control Engineering and Applied Informatics*, Vol.21, pp.3-11, 2019.

[4] A. K. Chandran, L. A. Poh, and P. Vadakkepat, "Real-time identification of pedestrian meeting and split events from surveillance videos using motion similarity and its applications," *Journal of Real-Time Image Processing*, Vol.16, pp.971-987, 2019.

[5] W. Sultani, C. Chen, and M. Shah, "Real-world anomaly detection in surveillance videos," *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6479-6488, 2018

[6] W. Li, V. Mahadevan, and N. Vasconcelos, "Anomaly detection and localization in crowded scenes," *TPAMI*, 2014.

[7]<http://www.multitel.be/image/researchdevelopment/research-projects/boss.php>.

[8] H. Rabiee, J. Haddadnia, H. Mousavi, M. Kalantarzadeh, M. Nabi, and V. Murino. "Novel dataset for fine-grained abnormal behavior understanding in crowd," *13th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, 2016.

[9] W. Sultani, C. Chen, and M. Shah, "Real-world anomaly detection in surveillance videos," *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6479-6488, 2018

### Acknowledgement

본 논문은 서울기술연구원 (19-4-5. CCTV 범죄상황 인지 기술 개발 연구)의 지원을 받아 수행된 연구임.