이상행동 감지 모델 시스템

System for Anomalous Behavior Detection

경북대학교 컴퓨터학부 김승준, 우승택, 박한솔, 김세아, 김민준, 정설영

(주)우경 정보기술 김호영

요약

최근 "묻지마 범죄"가 증가하고 있어 이러한 범죄를 줄이기 위한 노력이 확대되고 있다. 이상행동 감지 시스템은 두 가지 주요 방법을 사용하고 있다. 직접 관제인에 의한 수동 감시 방식과 인공지능을 활용한 이상행동 감지 방식이다. 직접 관제인을 통한 감시는 인력 소모가 크다는 단점이 있다. 하지만 AI 기반의 이상행동 감지는 일반적으로 공간 및 시간적 특성을 구별하여 객체와 그들의 움직임 정보를 추출하고, 공간적 및 시간적 특성을 결합하여 객체의 신원과 행동을 분류하기 때문에 인력 소모나, 경제적으로 효과적이다. 하지만 대부분의 인공지능 모델은 현실의 비디오의 상태를 고려하지 않고, 비디오의 모든 프레임을 동일하다고 가정하고 있다. 그러나 Low-Delay 모드와 같이 I-Frame과 P-Frame이 있는 경우, P-Frame은 일반적으로 I-Frame에 비해 정보가 적을 수 있으므로 이러한 가정은 모델이 학습을 불안전하게 할 가능성이 있다. 또한 QP(Quantization Parameter)가 높을수록 비디오 프레임의 상세 정보가 크게 손실되어 모델이 움직임과 행동을 정확하게 파악하는 데 어려움을 겪을 수 있다. 따라서 본 논문에서는 비디오 부호화 기술을 고려하여 인공지능 모델을 사용한 웹 기반 감지 시스템을 제안한다. 구체적으로는 인공지능 모델을 사용하여 관리자에게 실시간으로 알림을 보내고 해당 내용을 다중 서버 구조를 통해 처리해, 사용자가 웹 기반 UI를 통해 이상행동을 확인할 수 있는 시스템 모델을 설계하고자 한다.

서론

연구 배경

문제점 1 : 기존의 인공지능 모델들은 비디오 상태를 고려하지 않고, 모든 프레임을 동일하게 다룬다. 하지만 대부분 현실의 비디오는 압축된 상태로, 현재 많이 사용되는 기술인 H.264/AVC(Advanced Video Coding)와 H.265/HEVC(High Efficiency Video Coding)로 압축되어 있어, Low-Delay 와 Random-Access 모드 과정에서 프레임 정보 손실과 화질 변화로 인해 객체의 움직임과 행동을 정확하게 파악하기 어려워진다.

문제점 2 : 기존의 관제 시스템은 실시간으로 동영상 데이터를 처리 하는 특성 때문에 서버의 오버헤드가 크다.

연구 목적 및 제안 기법

관제 시스템의 특성과 비디오 부호화 기술의 특성을 고려하여, 인공지능 모델의 객체 탐지성능을 향상하고 효율적으로 영상 정보를 처리할 수 있는 자동 관제 시스템을 제안한다.

1. 입력 데이터인 비디오의 특성을 고려하여 모델의 성능을 향상.
2. 효율적인 다중 서버 구조를 사용하여 실시간 비디오 처리 속도 향상.

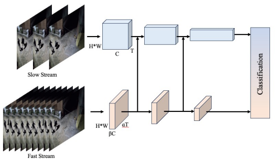
관련 연구

비디오 행동 인식

SlowFast Model : 비디오 행동 인식 딥러닝 모델

Slow Stream : 공간적 특성 파악, 큰 프레임 간격(τ), 채널 용량(C)

Fast Stream : 시간적 특성 파악, 짧은 프레임 간격(τ/α), 채널 용량(C*β)

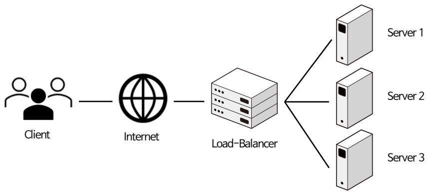


SlowFast Model	시간적 특성	공간적 특성	프레임 간격	채널 용량
Slow Stream	X	O	τ	C
Fast Stream	O	X	τ/α	C*β

분산 서버 구조

다중 서버 구조 : 두 개 이상의 서버를 연계·운용하는 기술로써 여러 개의 Transaction 요청이 들어왔을 때, 급격한 트래픽 변화에 대응하고 안정적으로 서비스를 제공하기 위한 기술.

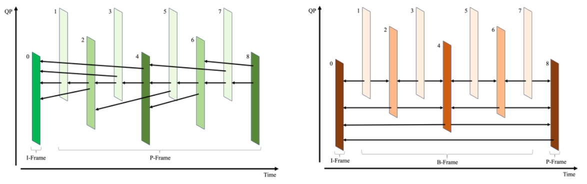
로드 밸런싱 : 외부의 요청을 서버에 적절히 분산하여 처리하는 기술.



비디오 부호화 기술

Low-Delay 모드 : Low-Delay 모드는 화상 통화, 스트리밍 방송, CCTV 등 다양한 분야에서 사용되는 비디오 부호화 모드. 프레임 외부의 정보로 부호화 하는 P-Frame(Inter 프레임 코딩)과 프레임 내부의 정보로 부호화 하는 I-Frame(Intra 프레임 코딩)을 사용.

Random-Access 모드 : 고화질 동영상 서비스 제공 등 다양한 분야에서 사용함. Low-Delay와 다르게 과거와 미래 프레임까지 활용하는 B-frame을 추가로 사용.



Low-Delay 모드　　Random-Access 모드

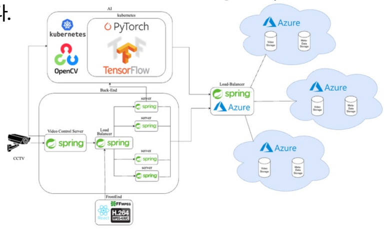
제안 모델

인공지능 모델

기존 SlowFast 모델은 비디오 내의 프레임 간 특성을 고려하지 않고, 일정 구간마다 프레임을 기반으로 학습한다. 특히, Low-Delay로 부호화된 비디오를 학습할 때, Slow Stream은 높은 QP를 가진 P-Frame을 선택하면 정보 손실이 크고 최적화가 어려운 결과를 초래할 수 있다. 이로 인해 추론 과정에서 원하는 결과가 나오지 않을 가능성이 있다. 또한, Fast Stream에서도 높은 QP를 가진 P-Frame을 선택할 경우, 프레임 내의 노이즈와 모션 블러로 인해 객체의 행동을 정확하게 표현하기 어려울 수 있다. Random-Access 또한 위와 같은 문제가 발생할 수 있다.

따라서 이러한 문제를 해결하기 위해 Slow Stream에는 정보 손실이 적은 I-Frame이나 낮은 QP를 가진 프레임을 우선적으로 사용하여 모델이 프레임 내의 객체를 명확히 인식하도록 하고, Fast Stream에도 정보 손실이 적은 프레임을 중점적으로 학습하여 객체의 행동을 정확하게 파악하도록 하는 것이다. 이를 통해 최적화된 결과에 도달하고 성능을 향상시킬 수 있을 것으로 예상된다.

시스템 아키텍처

현재 중앙 집중적인 아키텍처는 트래픽 변화에 대응이 어려우며, 멀티미디어 전송 비용이 높아 사용자에게 버퍼링과 같은 문제를 유발한다. 이러한 단점을 극복하기 위해 다중 서버 기반의 관제 시스템을 제안한다. CCTV로 생성된 영상은 비디오 제어 서버를 통해 기록되고 동시에 실시간으로 전달된다. 또한 인공지능 서버에서 이상행동 객체를 감지하고 해당 정보를 클라우드 기반 파일 시스템 서버에 저장한다. 클라이언트 서버는 실시간이나 과거 영상 요청 시 로드 밸런서를 통해 오버헤드가 적은 서버 주소를 받아와 연결을 형성한다.

이때 사용되는 스케줄링 기법은 Multi-Level Queue Scheduling으로, 파이프라인을 구축하면 클라이언트는 연결 종료까지 지속적으로 멀티미디어 데이터를 전송한다. 또한 클라우드에 저장된 데이터에 접근 시 클라우드 로드 밸런서를 사용하여 I/O 연산을 최소화하는 클라우드와 연결한다.