

AI 기반 데이터 마이닝 및 데이터 다이어트 VMS 시스템 제안

Proposal for an AI-based Data Mining and Data Diet VMS System

권민재¹ · 이재혁¹ · 김수정¹ · 김민성¹ · 정한솔¹ · 김병국^{1*}

¹삼육대학교 컴퓨터공학부

Min-jae Kwon¹ · Jae-hyeok Lee¹ · Su-Jeong Kim¹ · Min-Seong Kim¹ · Han-sol Jung¹ · Byoung-kug Kim^{1*}

¹Division of Computer Science and Engineering, Sahmyook University, Seoul, 01795, Korea

[Abstract]

This paper presents an AI-driven video data mining and data-diet system for CCTV monitoring. The system performs real-time object detection to produce metadata, assesses segment importance, and adaptively switches compression pipelines based on that assessment. High-importance segments are preserved at high resolution, while low-importance content is stored at reduced quality; the two classes are written to separate storage tiers to simplify operations and retention. By combining importance-aware analysis with dynamic encoding, the approach reduces overall storage footprint and accelerates retrieval focused on critical events. We detail the system architecture and control logic, outline an implementation on SBC-based edge devices with a VMS backend, and report results from a 24-hour deployment that demonstrate feasibility and benefits. The method operates without manual labeling, relying on object-count dynamics to guide decisions and quality shifts.

I. 서론

오늘날 사회는 로그, 이미지, 영상 등 대규모 비정형 데이터가 폭발적으로 생성되는 환경에 놓여 있다. 이러한 데이터에서 가치 있는 정보를 추출하고 패턴이나 규칙을 발견하는 데이터 마이닝(data mining) 기술은 다양한 분야에서 핵심 분석 도구로 자리 잡았다 [1]. 실제로 모바일 기기에서 발생하는 로그 데이터 분석, 이미지 검색 성능 향상을 위한 마이닝 연구 등이 활발히 진행되고 있다 [2],[3].

한편, 데이터 급증으로 저장 공간 소모가 심화되면서 효율성을 높이기 위한 데이터 다이어트(data diet)의 필요성도 꾸준히 제기되어 왔다. 예컨대 불필요한 데이터를 배제하거나, 영상에서 특정 객체가 포함된 구간만 추출하여 저장 효율을 개선하려는 연구가 보고되었다. 그러나 기존 방법은 전체 영상 데이터에 대한 일관된 최적화를 제공하지 못하거나, 단순 압축 수준에 머물러 성능 개선에 한계를 보였다 [4],[5].

특히 CCTV 기반 관제 시스템에서는 하루 종일 축적되는 방대한 영상 중 상당수가 정적 장면으로 구성되어 저장 공간을 불필요하게 점유하고, 관리자가 중요한 상황을 파악하기 어렵게 만든다. 이에 본 연구는 기존 CCTV VMS(Video Management System)에 AI 기반 데이터 마이닝과 가변 압축 방식을 결합한 새로운 영상 데이터 다이어트 시스템을 제안한다. 핵심 사건을 포함한 구간만 효율적으로 추출·저장하여 저장 공간 절감과 관리 효율 향상을 동시에 달성하고자 한다.

II. 제안 설계

본 연구의 시스템은 CCTV와 로컬 컴퓨팅 노드를 결합한 SBC(Single Board Computer) 기반 엣지 컴퓨터와 CCTV 영상 관리 시스템(VMS)으로 구성된다. 각 관제 구역에는 라즈베리파이와 카메라 모듈을 설치해 영상을 지속적으로 수집하고, 촬영 데이터는 동시에 로컬 컴퓨팅 노드와 중앙 VMS로 전송된다.

로컬 컴퓨팅 노드는 수신한 영상 프레임을 실시간 분석하여 객체 탐지를 수행한다. 탐지 결과는 객체 종류, 개수, 발생 시각 등의 정보로 변환되어 메타데이터 형태로 VMS에 전달된다.

VMS는 여러 엣지 장치로부터 영상과 메타데이터를 집계하여 구간별 중요도를 평가한다. 본 연구는 관제 구역 내 객체 수의 변화를 핵심 지표로 삼았다. 평상시보다 많은 인원이 탐지되면 잠재적 사건 가능성이 높다고 판단해 해당 구간을 고해상도로 우선 저장하고, 정적이거나 활동이 적은 구간은 저해상도로 압축해 저장 자원을 절감한다. 또한 시간대(근무/비근무)와 요일별 특성을 반영해 기준치를 조정함으로써 상황에 따라 보다 정밀하게 중요도를 판단한다.

이와 같은 설계를 통해 방대한 CCTV 영상 데이터 중 필요한 구간만 선별·저장함으로써 저장 자원을 효율적으로 활용하고, 사건·사고 발생 시 빠르고 신뢰성 있는 검색을 가능하게 한다.

III. 제안 구현

앞서 설명한 바와 같이, VMS는 메타데이터를 기반으로 중요 구간을 판단하여 데이터 마이닝과 데이터 다이어트를 수행한다. 이러한 로직은 영상 수신 스레드, 메타데이터 처리 스레드, 제어 스레드로 분리되어 동작한다. 영상 수신 스레드는 CCTV로부터

표 1. VMS 시스템 영상 수신 파이프라인

Pipeline Properties	FHD	SD
Frame	1920 x 1080	640 x 480
Overall Bitrate	7Mb/s	2.5Mb/s

영상을 받아 저장하며, 제어 신호에 따라 영상 수신 파이프라인을 재구성할 수 있다. 메타데이터 처리 스레드는 로컬 컴퓨팅 노드가 전달한 객체 탐지 결과를 데이터베이스에 기록하고 이를 제어 스레드에 전달한다. 제어 스레드는 전달받은 정보를 바탕으로 구간별 중요도를 판단하고, 결과에 따라 압축 옵션을 전환한다. 이 과정에서 VMS는 두 가지 파이프라인을 교대로 운용하여 데이터 다이어트를 수행하도록 설계되었다.

표 1과 같이, 두 파이프라인은 코덱(H.264)과 프레임 속도(30 FPS)는 동일하되 해상도와 비트레이트를 달리한다. 이를 통해 중요도가 높은 장면은 FHD로 저장해 정보 손실을 최소화하고, 중요도가 낮은 구간은 SD로 압축해 저장 자원을 절약한다.

IV. 실험 및 결과

제안한 시스템의 효과를 검증하기 위해 SD, FHD, 제안 시스템 세 가지 옵션을 각각 24시간 운용하여 생성된 총 영상 용량을 비교하였다. 용량 산정은 GiB 기준으로 수행하고 MiB 이하는 반올림하였다. 실험은 평일 09:00부터 익일 09:00까지 진행하였다. SD와 FHD 조건은 아래와 같다.

- SD 옵션: 640×480, 2.5 Mb/s
- FHD 옵션: 1920×1080, 7 Mb/s

실험 결과 그림 1과 같이, SD 24 GiB, FHD 72 GiB, 제안 시스템 33 GiB로 측정되었다. 제안 시스템은 FHD 대비 39 GiB 절감(54.17%)을 달성하여, 저장 효율 향상에 유의미하게 기여함을 확인하였다.

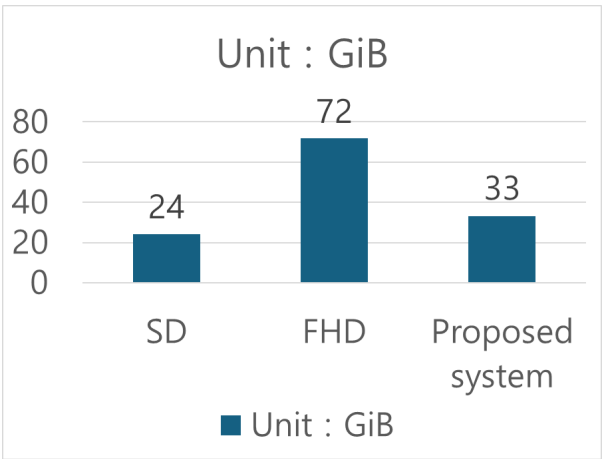


그림 1. 영상 데이터 용량 비교 차트

V. 결론

본 연구는 선별 로직에 데이터 다이어트를 결합한 AI 기반 데이터 마이닝 시스템을 제안하였다. 시간대·요일·근무시간 여부를 반영한 객체 수 변화를 지표로 삼아 중요 구간을 세밀하게 구분하고, 중요도가 낮은 구간은 저해상도로 저장하여 공간 사용을 줄이면서도 장면의 맥락과 시간적 연속성을 유지하도록 설계하였다. 동일 환경에서 24시간 운용한 결과, 제안 시스템은 FHD 대비 저장 용량을 54.17% 절감하는 효과를 보였다.

그러나 본 구현에는 몇 가지 한계가 존재한다. 영상의 중요도를 이분법적으로만 구분하여 세밀한 데이터 마이닝이 불가능했다. 또한 단일 공간에서 한 대의 옛지 컴퓨터로 평가를 수행해 일 변화 검증이 충분하지 못했다.

향후에는 넓은 공간에서 다수의 옛지 컴퓨터를 사용하여 충분한 검증을 수행하고, 다단계 중요도 분류를 도입하여 세밀한 데이터 마이닝이 가능하도록 개선할 계획이다.

Acknowledgments

본 연구는 2025년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업 지원을 받아 수행되었음(2021-0-01440)

References

[1] D. J. Park and W. S. Kim, “Improvement of the Parallel Importation Logistics Process Using Big Data” *The Journal of information and communication convergence engineering*, Vol. 17, No. 4, pp.267-273, Dec. 2019. DOI:<https://doi.org/10.6109/jicce.2019.17.4.267>.

[2] N. H. Cho, Y. G. Kang, “A Research Trends about Spatio-temporal Data Mining and Visualization of Log Data” *Journal of the Korean Cartographic Association*, Vol. 16, No. 3, pp. 15-27, 2016. Retrieved from <https://www.kci.go.kr/kciportal/ci/sereArticleSearch/ciSereArtiView.kci?sereArticleSearchBean.artiId=ART002185973>

[3] C. W. Lee, Y. M. Nah, “Design of Intelligent Image Retrieval System using Data Mining” *Korea Multimedia Society*, pp. 1-15, May. 1995. DOI:<https://dx.doi.org/10.6109/jkiice.2013.17.4.902>

[4] S. I. Kim, “A model to secure storage space for CCTV video files using YOLO v3” *Journal of The Korea Society of Computer and Information*, Vol. 28, No. 1, pp. 65-70, Jan. 2023. DOI:<http://doi.org/10.9708/jksoci.2023.28.01.065>.

[5] J. S. Seo, J. S. Shim, S. C. Park, “Design of Advanced City Support System through the CCTV Video Information BigData Analytics” *Korea Information Processing Society*, Vol. 21, No. 1, pp. 939-940, April. 2014. Retrieved from <https://korea-science.kr/article/CFKO201435553774102.pdf>