

실 전 문 제 연 구

CCTV영상에서 얼굴인식과 사람 재식별을 이용한 확진자 고속 검색 시스템

팀 명 마지막 빌드
발 표 자 김 도 희

Parlab

(병렬소프트웨어설계연구실_이영민교수님)

CONTENTS

01

연구 소개

02

배경 지식

03

시스템 개요

04

가속화 기법

05

연구 결과

06

결론 및 한계

01

연구 소개

01

역학조사 방법론 1 – 휴대전화&카드 사용 내역 기반



통신사 기지국 기반 위치정보 & 카드 사용 정보

-> 단편적 정보 제공!

역학조사 방법론 2 – 인터뷰 기반



방법론 1과 확진자와의 인터뷰를 통해 동선 확정

-> 확진자의 거짓진술 & 연락두절 문제

-> CCTV 확인을 통해 최종 동선 확정(수작업)

제안하는 시스템

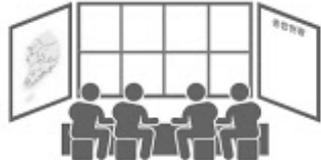
“ 확진자의 동선을 자동으로 찾아주는 실시간 **지능형 CCTV 검색 시스템**”

1. 딥러닝 기반 얼굴 검출
2. 얼굴인식
3. 객체 검출
4. 사람 재식별 기술

네 가지의 기술을 활용하여 확진자 동선을 추적하는 시스템 설계

01

구체적 동선 제공



CCTV 관제(수작업)

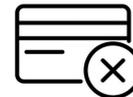


수작업

자동화



확진자와의 인터뷰



CREDIT CARD

휴대전화&신용카드 사용 정보

단편적 동선 제공

Our contribution

1. 얼굴 이미지만으로 해당 인물 등장 동영상 추출

⇒ 사람 재식별에 필요한 훈련 데이터, 검색 데이터 생성

2. open-set 데이터셋 구축

⇒ 통제되지 않은 상황을 가정한 Full Frame 동영상 데이터셋 구축

3. 고속 검색을 위해 멀티 쓰레딩, 파이프라이닝 기법 적용

⇒ 동시에 여러 채널의 동영상에서 고속의 사람 재식별 기능 수행

02

배경 지식

사람 재식별 (Person Re-ID)

“여러 CCTV 동영상에서 동일한 사람을 탐지 및 매칭하는 기술”



특징

1. 얼굴 인식이 불가능한 경우에도 보행자의 외형을 바탕으로 매칭 가능
2. CCTV 정보, 시간이 저장되어 특정 인물의 동선파악에 유리함
3. 값비싼 사람의 노동력을 대체할 수 있음.

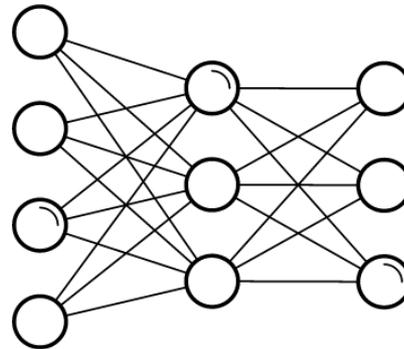
사람 재식별 과정

1. 입력 이미지로부터 특징 벡터 추출

입력 사진



사람 재식별 모델



추출된 특징 벡터

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_{n-1} \\ X_n \end{bmatrix}$$

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_{n-1} \\ Y_n \end{bmatrix}$$

Pedestrian Images

Re-id model

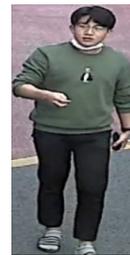
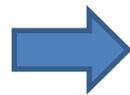
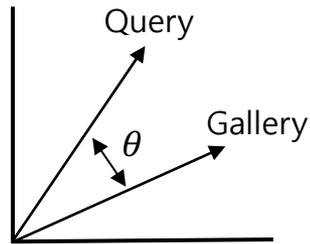
Feature extract

사람 재식별 과정

2. 추출된 특징 벡터 사이의 유사도 비교

Similarity(A, B)

Similarity(Query, Gallery) > τ



Query





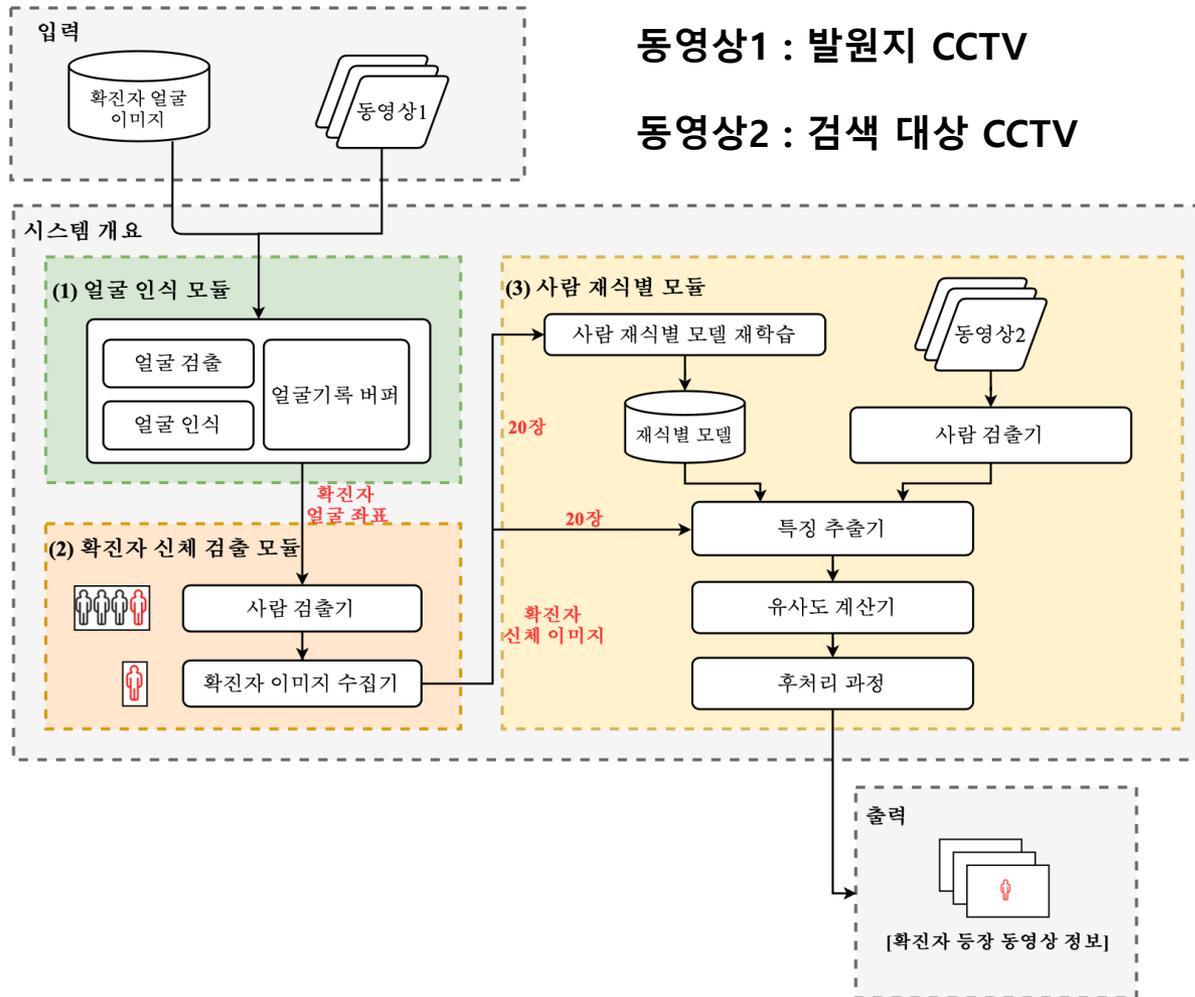

$$\frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}}$$



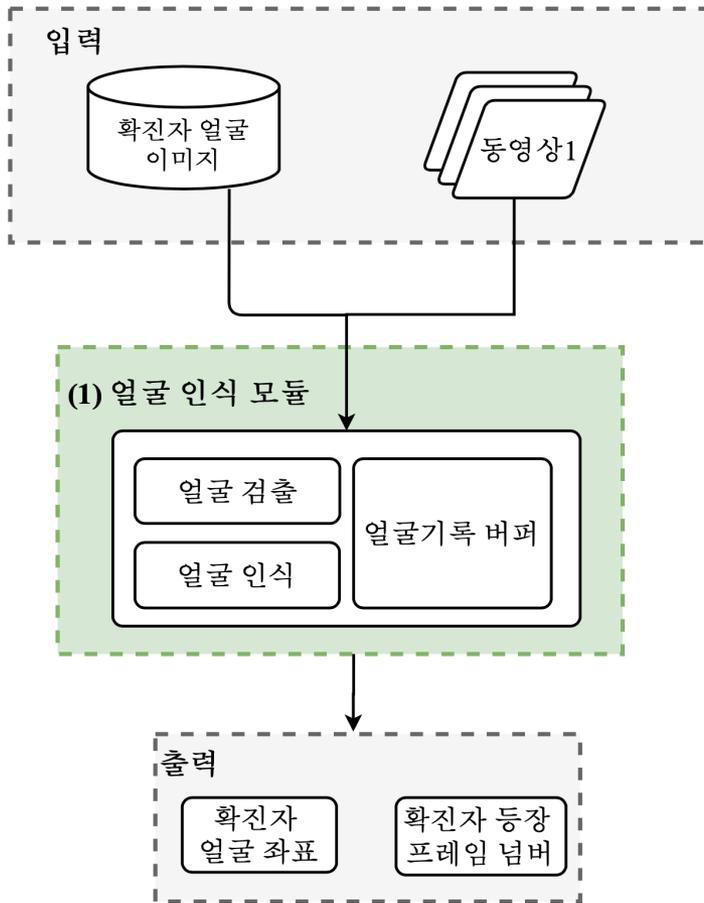
03

시스템 개요

시스템 개요



얼굴 인식 모듈

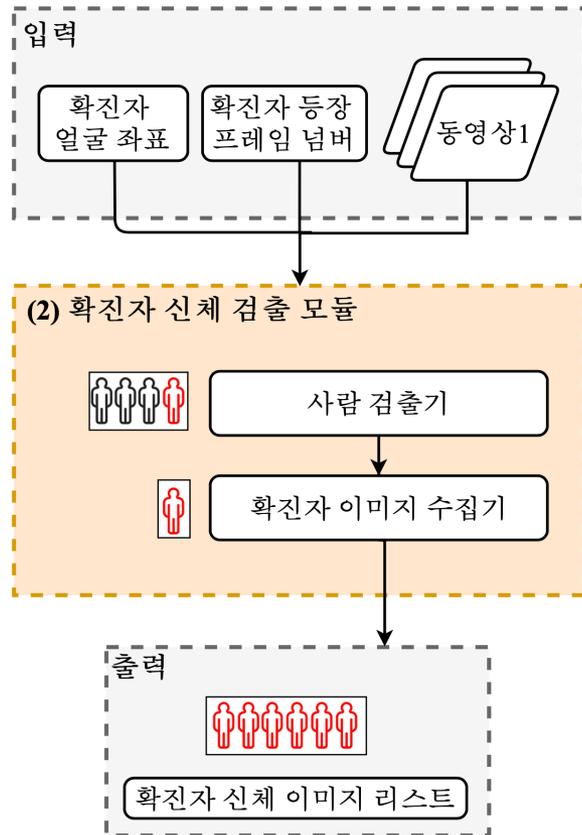


- 확진자의 얼굴 + 발원지 동영상을 입력으로 받음
- 얼굴 검출엔 RetinaFace를, 얼굴 인식엔 ArcFace를 사용
- 동영상에서 얼굴 인식 정확도 향상을 위해 본 연구실에서 제안한 얼굴 기록 버퍼 사용
- 얼굴 기록 버퍼?
연속된 프레임에서 비슷한 위치에 동일한 인물이 나온다는 가정으로 과거 얼굴 인식 정보를 토대로 얼굴을 인식

Deng, J., Guo, J., Zhou, Y., Yu, J., Kotsia, I., & Zafeiriou, S. (2019). Retinaface: Single-stage dense face localisation in the wild. *arXiv preprint arXiv:1905.00641*.
 Deng, J., Guo, J., Xue, N., & Zafeiriou, S. (2019). Arcface: Additive angular margin loss for deep face recognition. *CVPR* (pp. 4690-4699).
 최재권, 김도희, 박경철, & 이영민. (2019). 얼굴 객체 및 휴리스틱 알고리즘을 활용한 비디오 실시간 얼굴 인식 시스템. *한국정보과학회 학술발표논문집*, 1749-1751.

03

확진자 신체 검출 모듈



확진자
이미지
수집기

확진자 신체 이미지 리스트

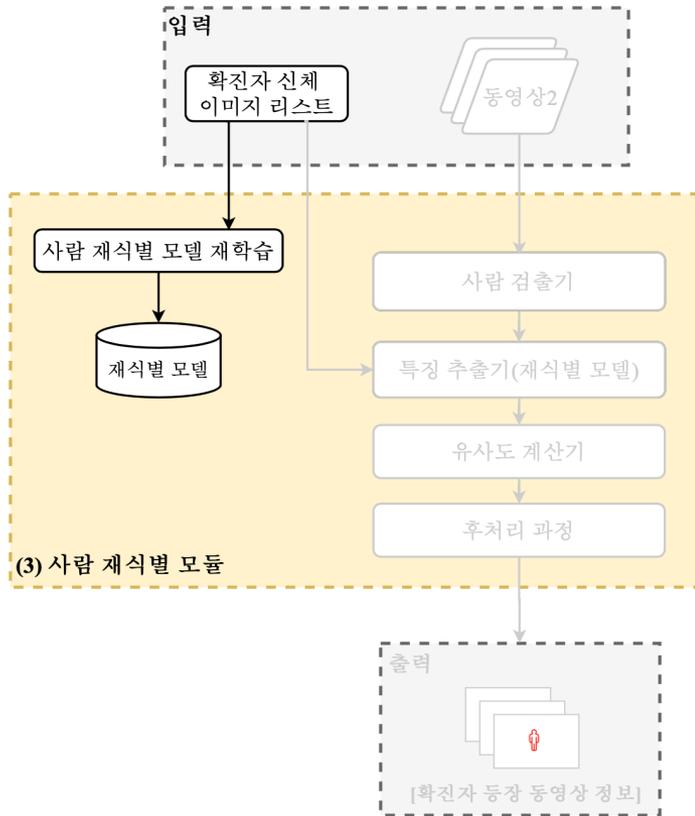


- 확진자의 얼굴 + 발원지 동영상을 입력으로 받음
- 사람 검출기에는 Yolo v4를 사용
- 얼굴 좌표와 사람 검출기에서 구한 신체 좌표간 IoU(Intersection over Union)을 계산해 확진자 신체 이미지를 검출 및 수집

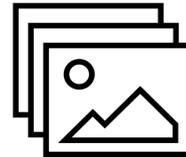
Bochkovskiy, A et al. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. *arXiv preprint arXiv:2004.10934*.

03

사람 재식별 모듈 - 재학습

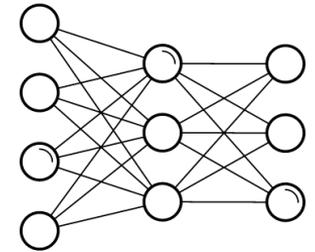


확진자 신체 이미지



Market-1501 데이터셋

Re-train
(Transfer learning)

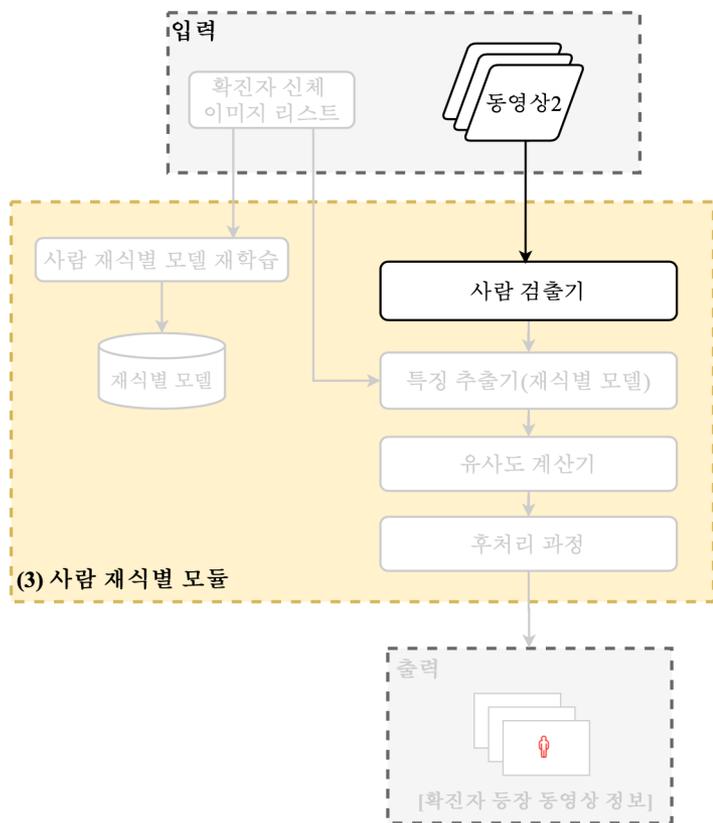


재학습 모델

- 확진자 신체 이미지(20장)을 Market 데이터셋과 합쳐서 새로운 학습 데이터셋을 구축
- 사람 재식별 모델을 재학습(transfer learning)시킴
- 재학습(소요시간:40분) 후에는 확진자의 신체 특징이 더 반영된 모델이 됨

Zheng, L., Shen, L., Tian, L., Wang, S., Wang, J., & Tian, Q. (2015). Scalable person re-identification: A benchmark. In *Proceedings of the IEEE international conference on computer vision* (pp. 1116-1124).

사람 재식별 모듈 - 사람 검출기

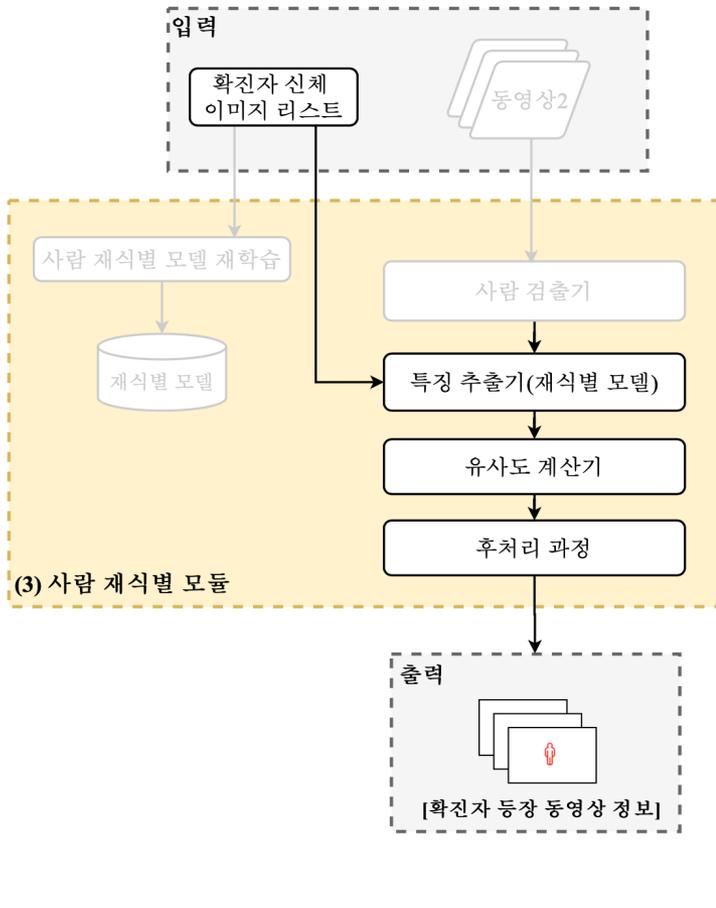


- CCTV에 등장하는 모든 보행자를 검출
- 검출된 보행자 신체 이미지를 사람 재식별 모델로 보냄

Bochkovskiy, A et al. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. *arXiv preprint arXiv:2004.10934*.

03

사람 재식별 모듈 - 특징 추출 및 유사도 계산

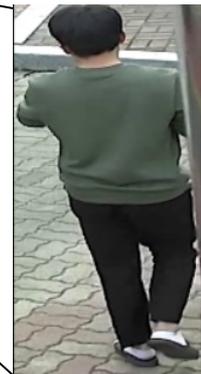
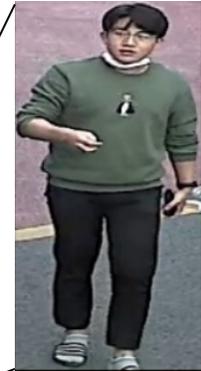
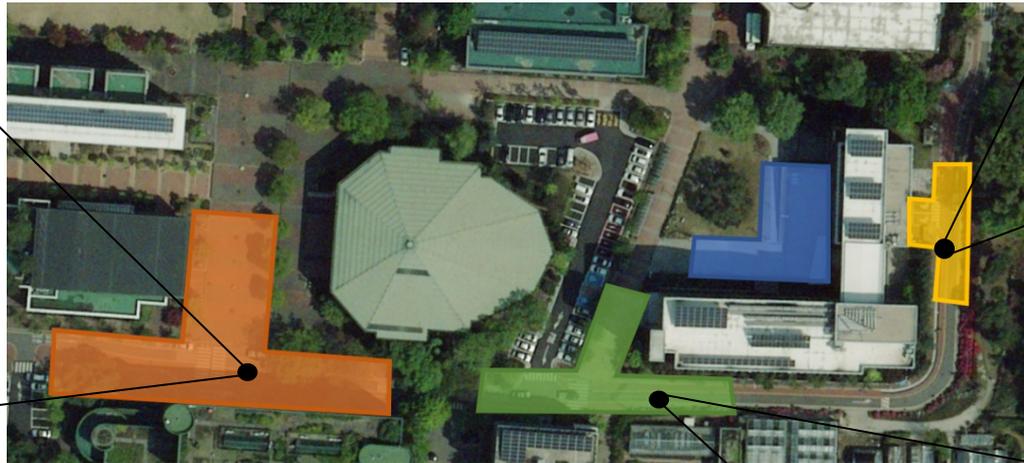


특징 벡터	유사도	정답
	1.0	
	0.9	
	0.2	
	0.1	

Zheng, Z., Zheng, L., & Yang, Y. (2017). A discriminatively learned cnn embedding for person reidentification. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 14(1), 1-20.

03

사람 재식별 모듈 - 확진자가 등장한 위치, 시간 정보 출력



출력 결과

C20201014.172500-20201014.180146_CH2/

⇒ CH2번 카메라에서 10월 14일 17시 25분부터 찍힌 영상

34712_03_0.977.jpg

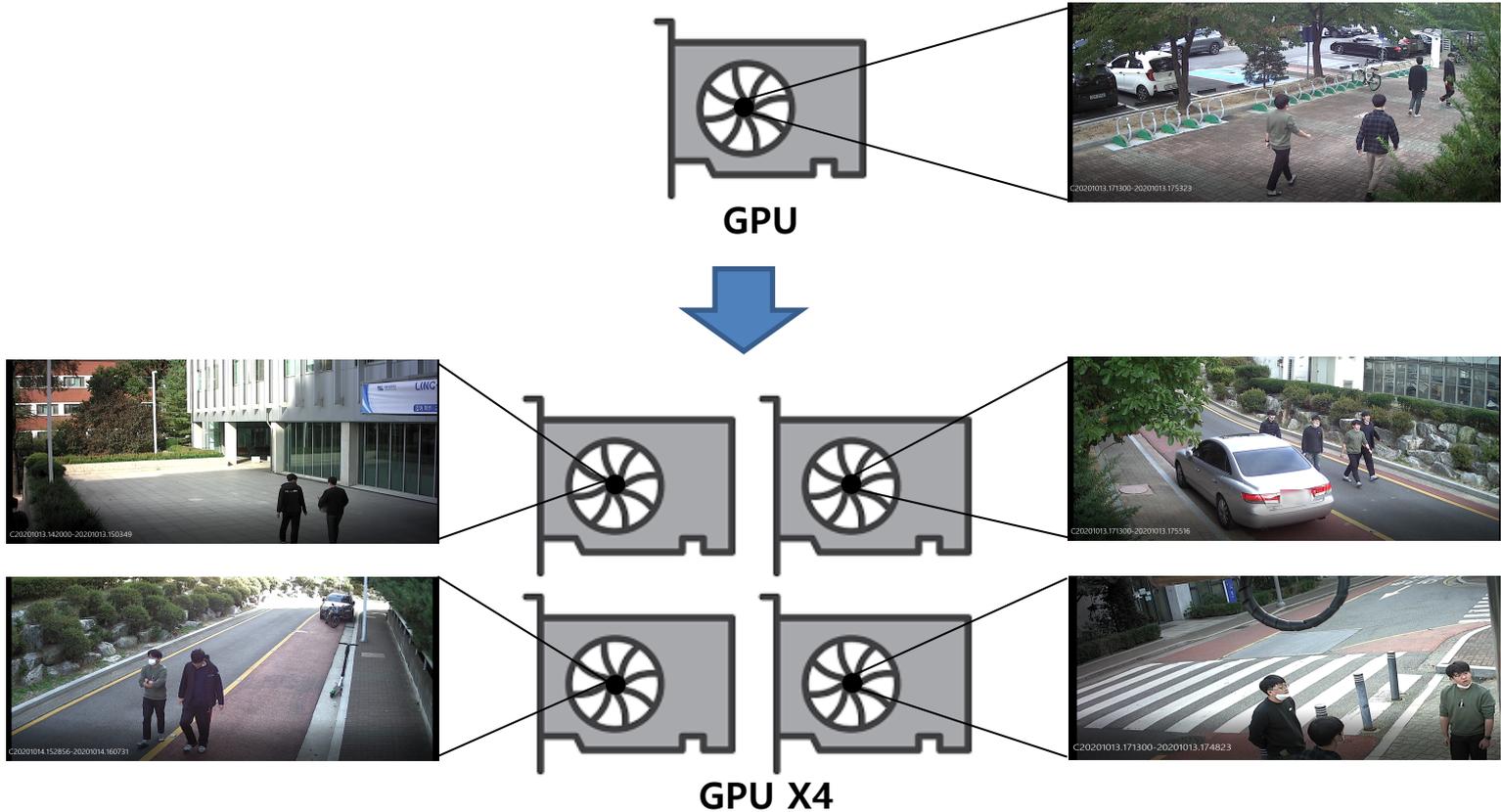
⇒ 34712번 프레임에서 검출된 사람은 확진자와 유사도가 0.977이다.

04

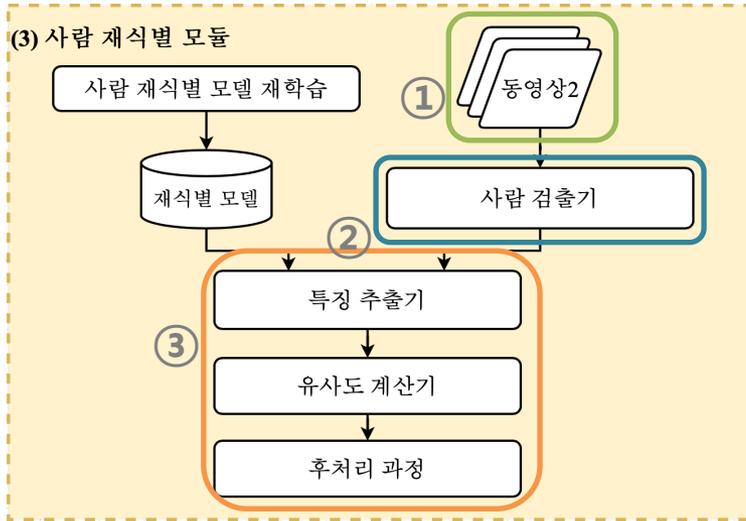
가속화 기법

멀티 쓰레딩

- GPU 1대가 하나의 동영상을 처리하도록 구현
- Multi-Threading을 통해 Multi-Channel 추론 가능



파이프라이닝



사람 재식별 모듈의 단계

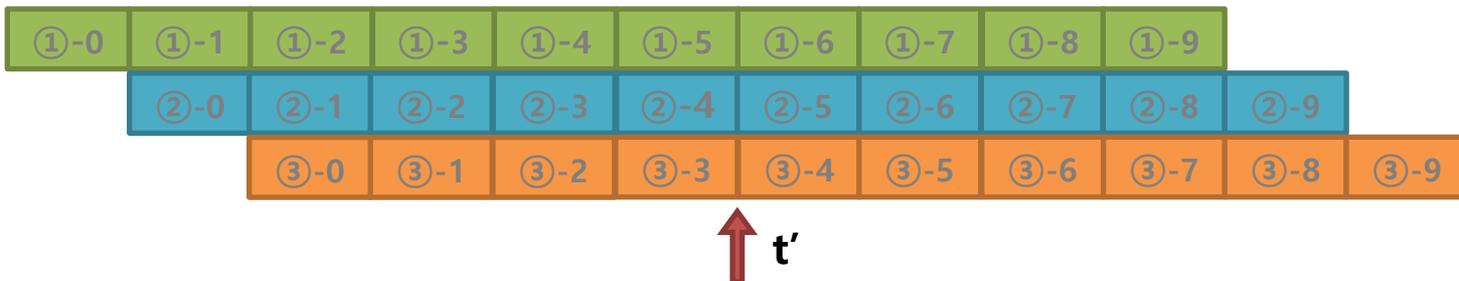
- ① 입력 동영상을 프레임으로 디코딩
- ② 사람 검출기
- ③ 사람 재식별 특징 추출기 + 유사도 계산 + 후처리

“각 단계를 파이프라이닝하여 병렬적으로 수행”

기존



파이프라이닝



05

연구 결과

데이터셋 구성

1) 기존 데이터셋의 한계

- 기존 데이터셋(Duke-MTMC, Market-1501)의 특징
 - 영상 형태의 자료가 아닌 사진으로 제공
 - 전체 이미지가 아닌 사람 부분만 잘라서 제공
- 우리가 제안하는 시스템은 CCTV 영상에서 확진자 고속 검색 시스템이므로, 학습에는 기존 데이터셋을 사용할 수 있으나 시스템 테스트에는 한계점을 보임.



Duke-MTMC Dataset Sample



Market-1501 Dataset Sample

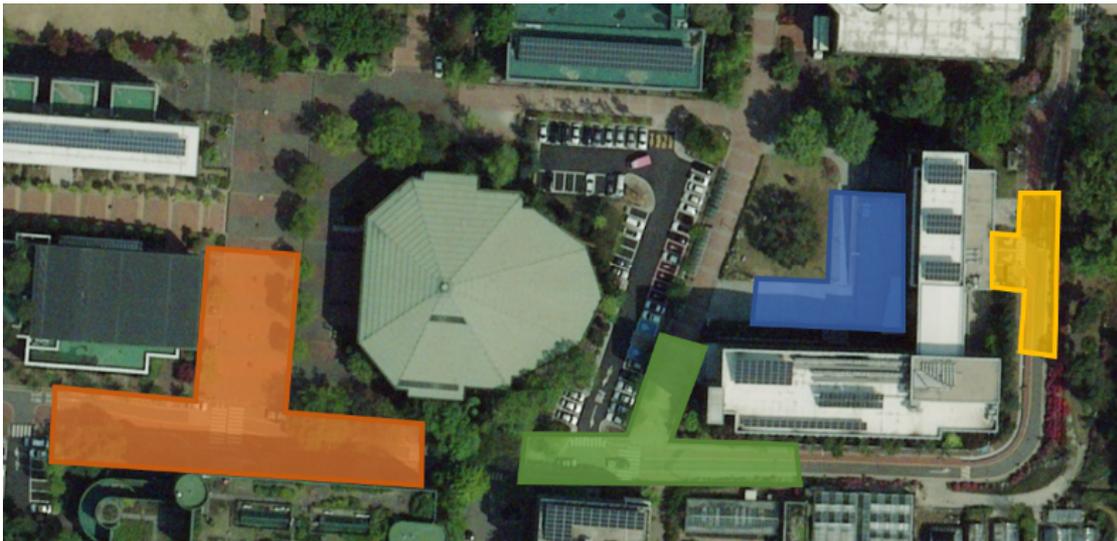
Gou, M., Karanam, S., Liu, W., Camps, O., & Radke, R. J. (2017). DukeMTMC4ReID: A large-scale multi-camera person re-identification dataset. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (pp. 10-19).

Zheng, L., Shen, L., Tian, L., Wang, S., Wang, J., & Tian, Q. (2015). Scalable person re-identification: A benchmark. In *Proceedings of the IEEE international conference on computer vision* (pp. 1116-1124).

데이터셋 구성

2) 데이터셋 제작

- 서울시립대학교 캠퍼스 4개 구역에서 촬영
- 한 구역당 3대의 CCTV. 높이 3m
- 촬영 시각은 조도의 영향을 알아보기 위해 낮 14~16시, 일몰 17~19시로 나누어 측정
- 총 CCTV 누적시간 19시간 30분. 확진자 대역 52분 48초간 등장



실험 결과

1) 정확도 평가



[CCTV 동영상 내 확진자 예시]

TTR(True Target Rate) 측정방법을 이용

TTR = 92.44%

 : 등록된 확진자 신체 이미지

 : True Positive

 : False Positive



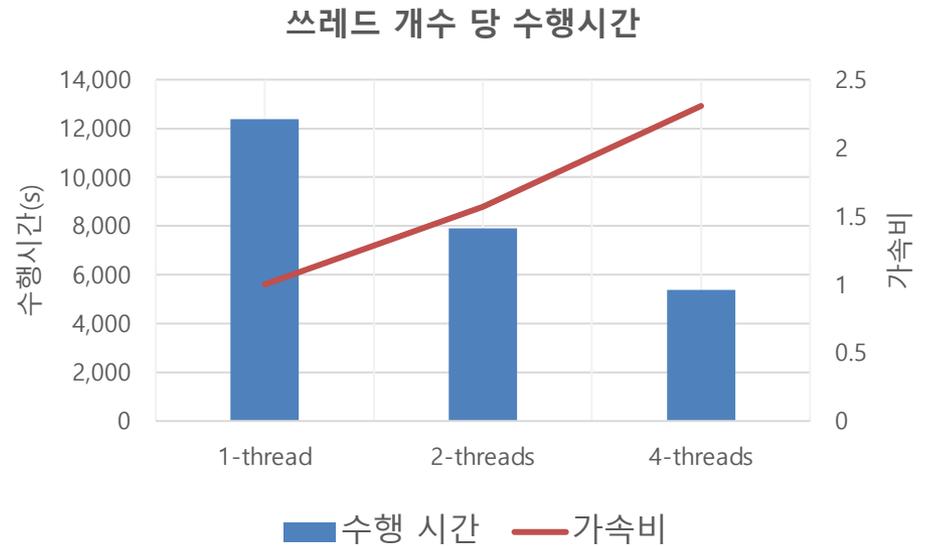
[사람 재식별 모듈 결과]

실험 결과

2) 효율성 평가

1. 스레드 개수에 따른 수행시간 평가

	1-thread	2-threads	4-threads
수행 시간	12,387초	7,893초	5,369초

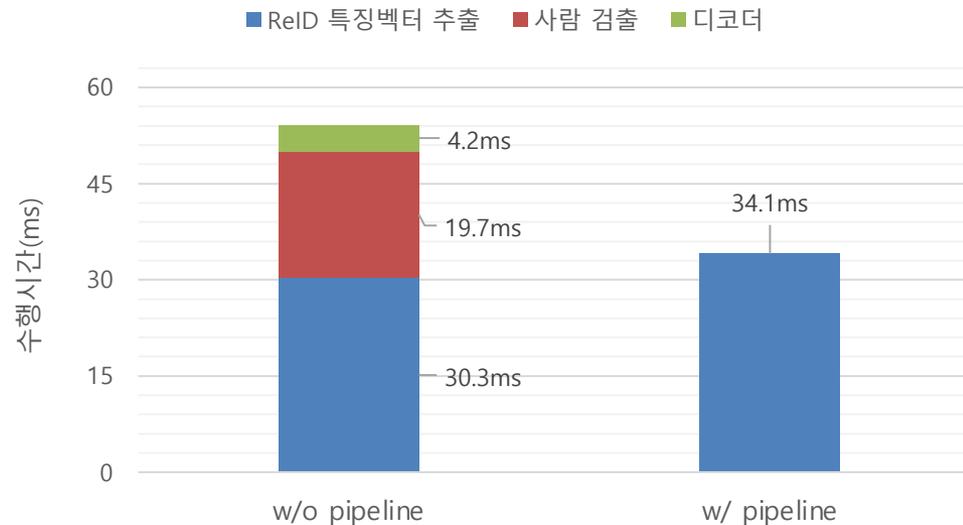


- 사람이 등장해야 특징 추출이 수행됨
- 공정한 측정을 위해 사람이 전체 중 93.2%의 프레임에서 등장하는 38분 동영상 4개 (누적 152분 = 9,120초)에 대해 실험
- 선형적으로 가속되지 않는 이유는 파이썬 GIL(Global Interpreter Lock)때문으로 보임
- 스레드 4개 사용시 스레드 1개에 비해 2.6배 가속됨

실험 결과

2) 효율성 평가

2. 파이프라이닝 유무에 따른 수행시간 평가



- 사람 재식별 모듈의 ①디코더 ②사람 검출 ③Re-ID 특징 추출 단계들을 파이프라이닝 수행
- 각 단계의 시간은 프레임당 ① 4.2ms, ② 19.7ms, ③ 30.3ms
- 파이프라이닝을 적용한 시간은 34.1ms 적용 전에 비해 1.6배 가속

06

결론 및 한계

결론

1. CCTV 동영상에서 특정 인물이 등장하는 영상을 추출하는 시스템 구축

⇒ 얼굴 인식과 사람 재식별 기술을 사용해 자동화된 확진자 검출 및 매칭 시스템 구축

⇒ TTR(True Target Rate) = 92.44 %로 높은 정확도를 얻음

2. 다양한 가속 기법을 활용한 시스템의 고속화

⇒ 멀티쓰레딩, 파이프라이닝 기법을 통해 사람 재식별 모듈을 3.7배 가속

⇒ FHD 해상도의 19.5시간 분량 동영상들을 5.4시간만에 자동 검색

한계

1. IR 카메라(infrared camera) 영상에서는 사람 재식별을 수행할 수 없음

=> 사람 재식별은 옷 색을 중요한 특징으로 보기 때문에 야간에 촬영되는 IR 카메라(infrared camera) 영상에서는 현재 시스템으로 사람 재식별을 수행할 수 없음



[IR 카메라 영상]

2. 마스크 착용 등 얼굴인식이 안되는 상황에서는 수작업으로 확진자의 신체 이미지를 입력 해야함.

THANK
YOU

발 표 자 김 도 희