

---

# 캡스톤 디자인 프로젝트 계획서

---

## 데이터 크롤링과 얼굴인식을 이용한 개인정보 영상 검출 시스템

이충헌(151779)

김태준(164160)

김병수(181526)



캡스톤 지도교수 서명 유 석 봉 

## [제목/주제]

데이터 크롤링과 얼굴인식을 이용한 디지털 개인정보(성범죄) 영상 검출 시스템.

## [배경 및 문제]

- 디지털성범죄 영상물 유통사례 적발 4,584 건 (2018.7.31. 방송통신위원회)
- 카메라 등을 이용 불법촬영 범죄 꾸준한 증가 (2019.12.30. 여성가족부)  
(‘12) 2,400 건 → (‘13) 4,823 건 → (‘14) 6,623 건 → (‘15) 7,623 건 → (‘16) 5,185 건 → (‘17) 6,470 건
- 피해를 최소화하기 위해 영상의 유출을 막기위한 빠른 조치의 필요성이 대두됨.
- 피해자가 자신의 영상이 어디에 유출되었는지 직접 찾기 어려움.

### <기대 효과>

- 1 ) 리벤지 포르노(디지털성범죄)영상 노출 이후에 빠른 후속조치가 이루어 질 수 있도록 커뮤니티, SNS 에서 최대한 빠르게 본인의 얼굴 등이 노출되었다고 의심되는 영상, 사진 등을 찾을 수 있음.
- 2 ) 유출에 따른 피해를 최소화하는 시스템을 개발함으로써, 개인정보의 무단 활용을 방지함.
- 3 ) 궁극적으로 데이터 수집을 통한 빠른 후속 대처 및 법적 대응에 용이하게 함.

## [목표]

영상 내 얼굴과 같은 개인정보를 인식함으로써, 개인정보 유출로 인한 피해 방지 기술 연구/개발

## [팀 구성]

<소프트웨어 공학과>

팀명 : 811 호

팀 리더 : 이충헌(151779)

팀원: 김태준(164160), 김병수(181526)

## [요구 사항]

요구사항 ID	명칭	설명
R001	데이터 수집	특정 사이트 크롤링.
R002	데이터 저장	크롤링된 데이터 저장.
R003	얼굴인식	얼굴인식 딥러닝 학습.
R004	매칭 시스템	피해자의 데이터 매칭 및 비교
R005	웹 클라이언트	키워드와 사진으로 검색 등 종합적으로 사용할 수 있는 시스템

요구사항 ID		R001	요구사항 명칭	데이터 수집
상세 설명	정의	사진 및 영상 데이터 수집		
	세부	크롤링을 통한 사진 및 영상 데이터 수집		
	내용	-크롤링 기법을 이용하여 개인정보 영상이 유출 되었을 거라고 예상되는 사이트에서 데이터 수집.		

요구사항 ID		R002	요구사항 명칭	데이터 저장
상세 설명	정의	데이터 저장		
	세부	크롤링 된 데이터를 서버에 저장		
	내용	-크롤링 된 데이터의 URL 을 DB 에 저장.		

요구사항 ID		R003	요구사항 명칭	얼굴인식
상세 설명	정의	얼굴 인식을 위한 딥러닝 학습.		
	세부	얼굴인식 모델 학습		
	내용	-한국인 얼굴 데이터셋을 이용하여 모델을 학습시킴.		

요구사항 ID		R004	요구사항 명칭	매칭 시스템
상세 설명	정의	피해자 이미지 매칭.		
	세부	데이터와 입력 이미지 비교.		
	내용	-학습된 AI 를 이용하여 피해자의 이미지를 수집한 데이터와 비교		

요구사항 ID		R005	요구사항 명칭	웹 클라이언트
상세 설명	정의	웹 클라이언트		
	세부	사용자가 쉽게 접근할 수 있는 UI 개발.		
	내용	-웹 서버 기반 클라이언트 개발		

## [간트 차트]

