



YouHi

이태훈 이주형 김성수 김민재 이인평



INDEX

01. 피드백 반영

- 1). 자막 욕설 검열 기능 추가 및 임베딩 기법 실험
- 2). 음성 마스킹 구현

02. 변경 및 추가 사항

- 1). 영상 검열
- 2). 서버 Instance 추가

03. 진행 현황

- 1). 영상 검열
- 2). 음성 및 자막 검열
- 3). Front-end
- 4). Back-end

04. 애로사항 및 기술적 한계



자막 욕설 검열 기능 추가 및 임베딩 기법 실험

자막 추출 및 다양한 임베딩 기법 실험

자막 추출: Tesseract-OCR, Google Vision API OCR
임베딩: Sentence Embedding(HanBert), Glove 실험

Tesseract-OCR

무료 서비스이지만 기능이 제한적이고, 정확도가 떨어짐

Sentence Embedding

최신 기술인 Google의 Bert 한국어 버전 HanBert로 실험
정확한욕설 필터와는 방향이 다르다고 판단

Google Vision API OCR

유료 서비스이지만 가격이 저렴하고,
정확도가 높고 강력한 기능을 제공

Glove

OOV (Out Of Vocabulary) 대처 불가
FastText보다 낮은 정확도



음성 마스킹 기능 구현

음성 마스킹 기능 구현

욕설이 탐색된 구간의 음성을 추출하고,
다른 음성(Beep Sound)으로 덮어씌워서 구현

욕설 탐색

Google STT에서 변환된 텍스트의 시간을 제공하는
기능으로 욕설의 정확한 위치를 탐색

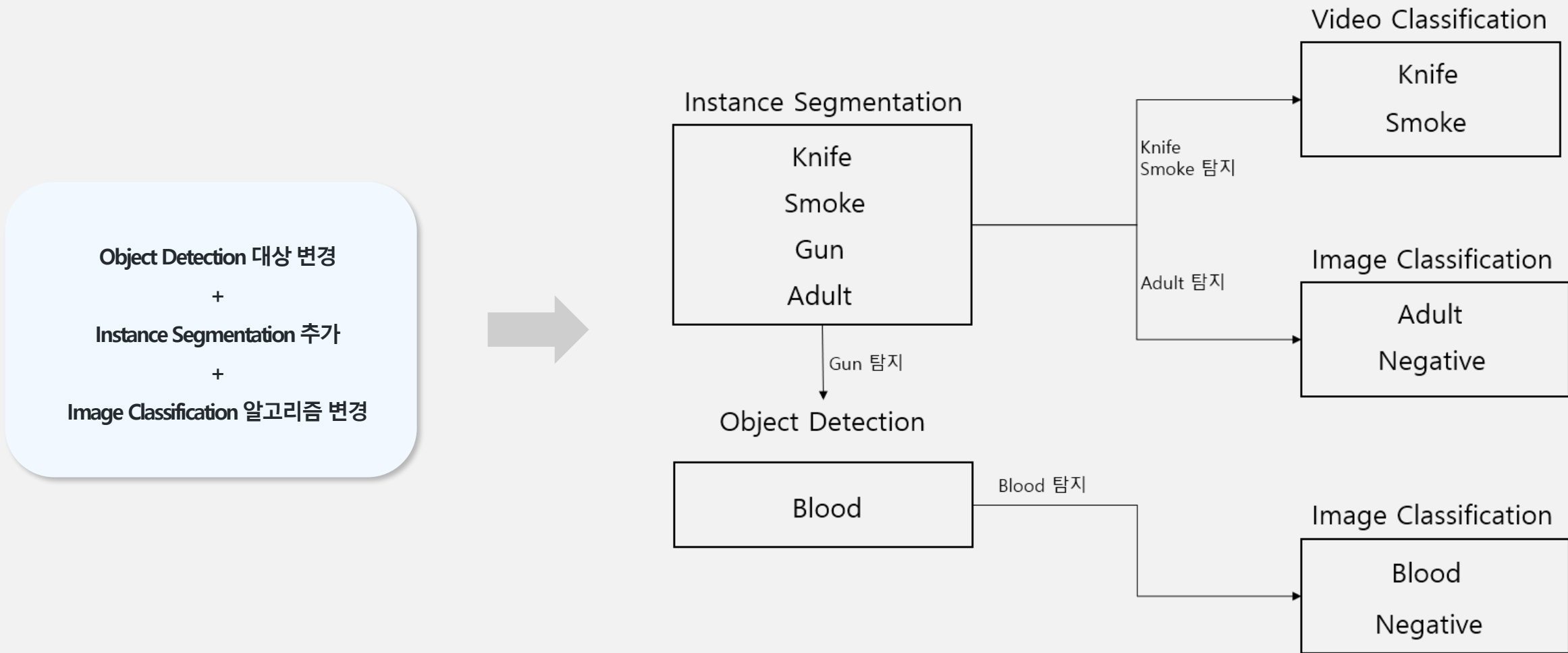
Original audio



Beep sound



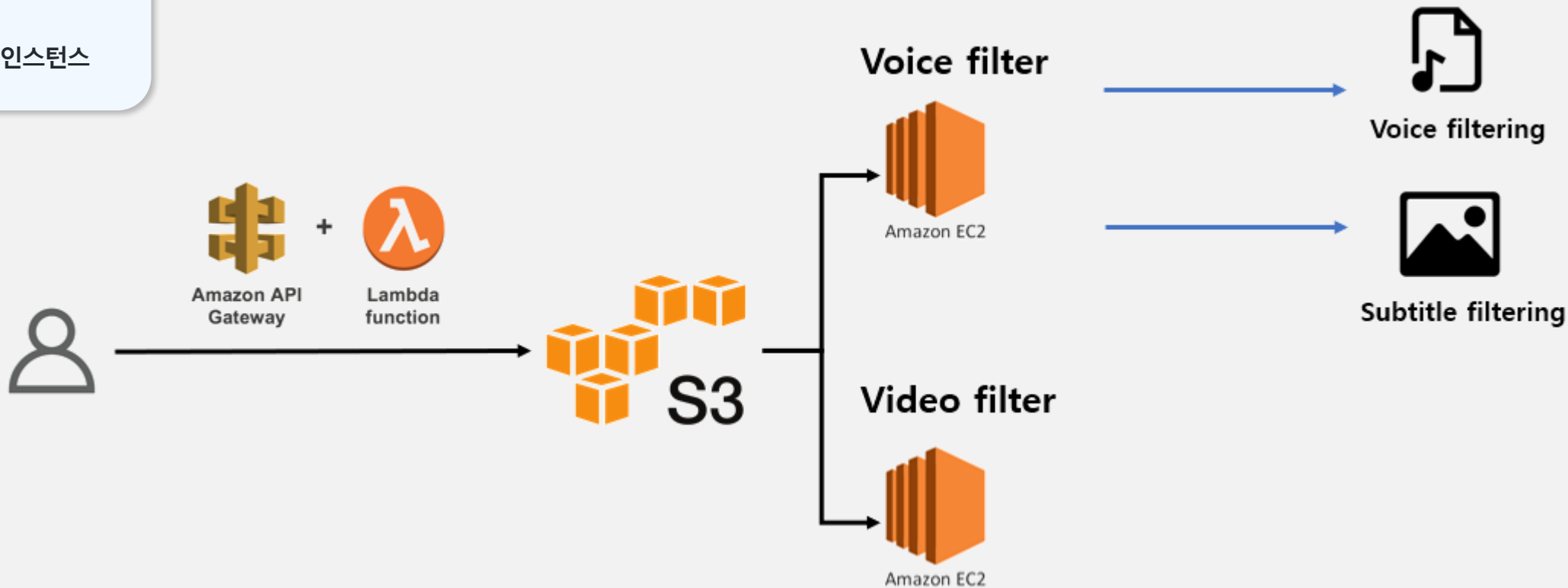
영상 검열





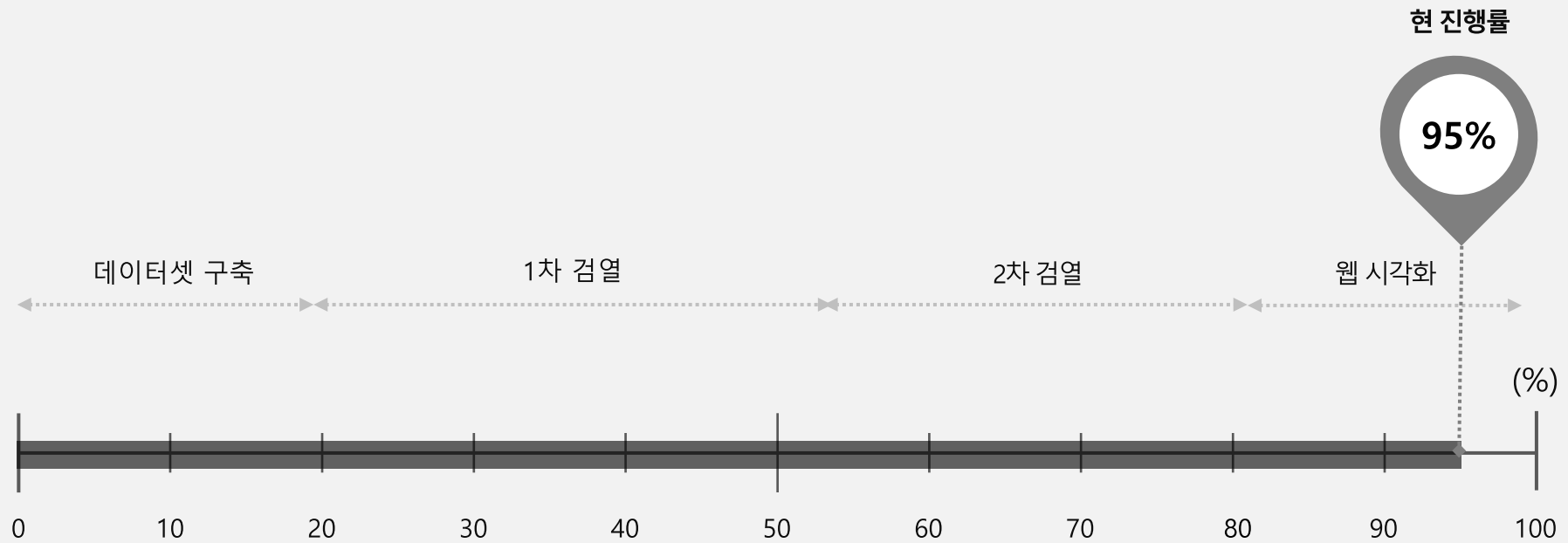
서버 Instance 추가

영상 검열 및 웹 서버 인스턴스
+
음성 및 자막 검열 인스턴스





영상 검열 진행도





Instance Segmentation - 데이터셋 구축

knife



smoke



gun



adult

...



70장씩



Labelme를 이용한 라벨링



Instance Segmentation - 결과



Knife



Smoke



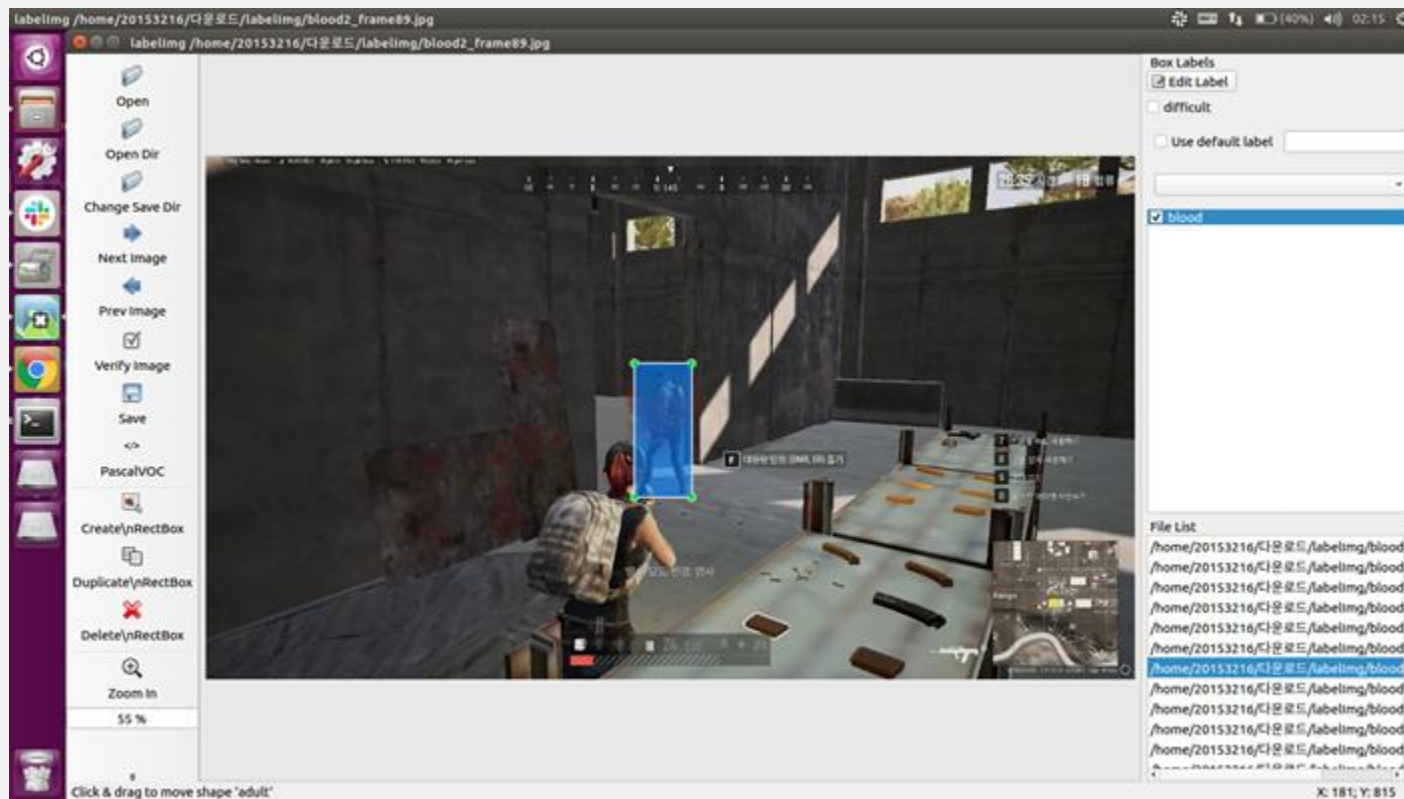
Gun



Adult



Object Detection - 데이터셋 구축



LabelImg를 이용한 라벨링



Object Detection - 결과



오분류(빨간 옷)



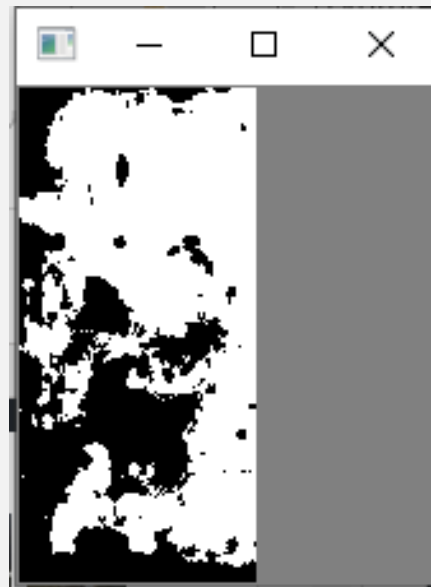
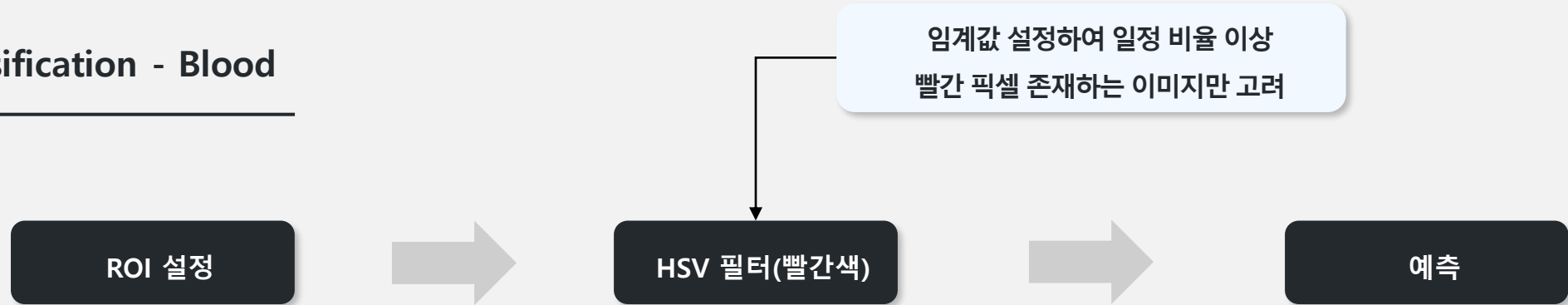
오분류(일반 캐릭터)



정분류



Image Classification - Blood



피를 흘리고 있는 사람인지
빨간 옷을 입은 사람인지 판단



Image Classification - Adult

ROI 설정



예측

신체의 상반신으로 ROI 설정

옷을 벗고있는지 입고있는지 판단



Video Classification – Knife, Smoke

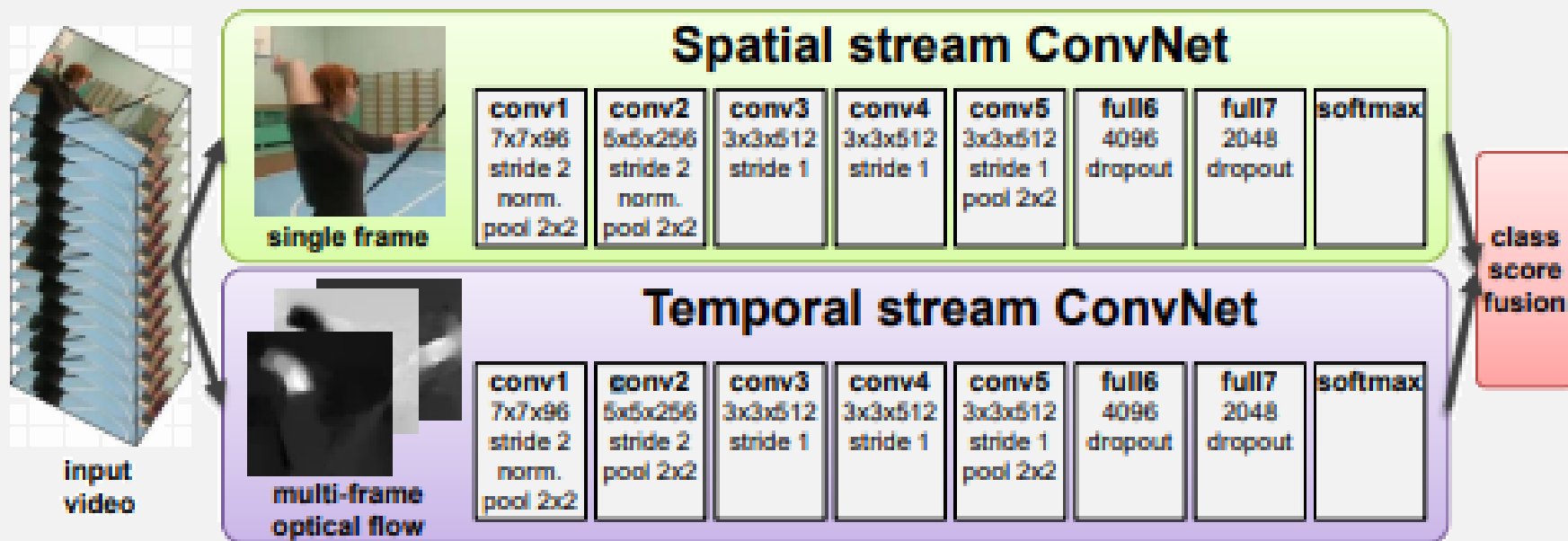


Figure 1: Two-stream architecture for video classification.



Video Classification – Knife, Smoke

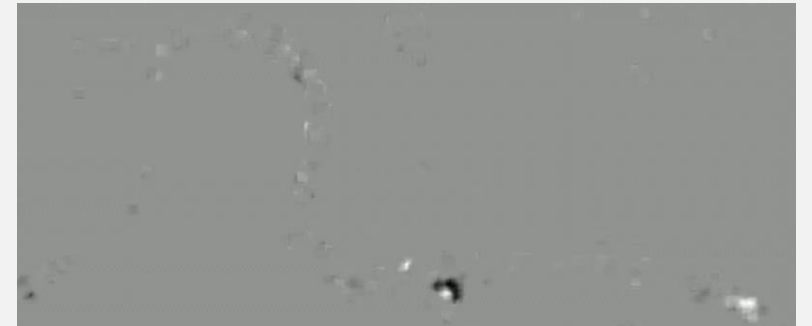
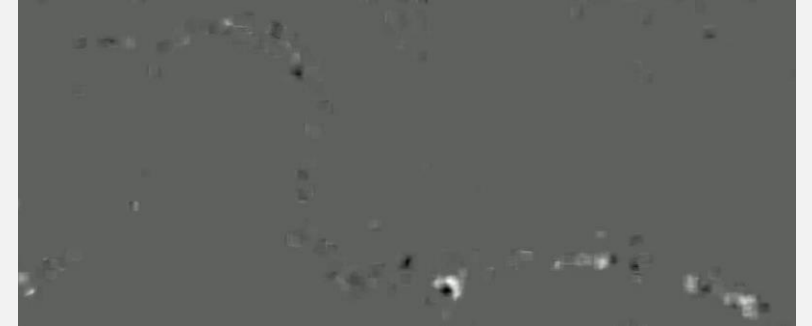


1	smoke4.mp4
2	0.912588
3	0.980637
4	0.993374
5	0.99166
6	0.996229
7	0.993992
8	0.986162
9	----
10	

spatial stream

1	smoke4.mp4
2	0.208731
3	0.391649
4	0.16849
5	0.98884
6	0.930769
7	0.981653
8	----
9	

temporal stream



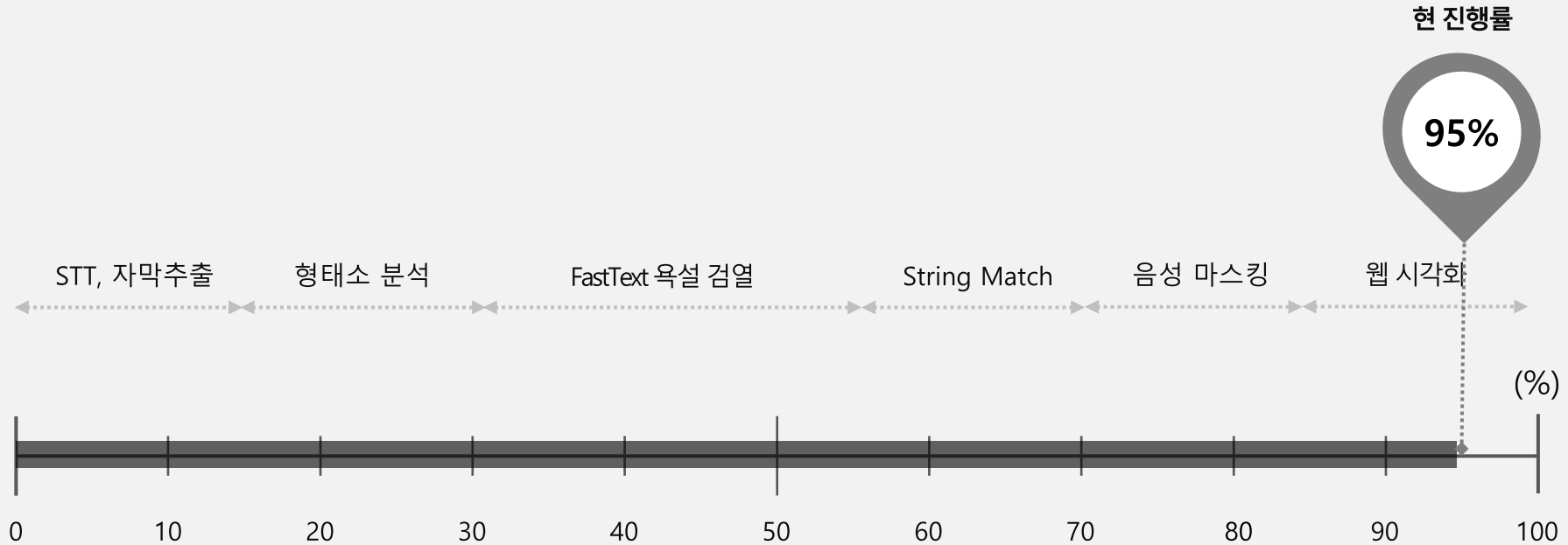
1	smoke4.mp4
2	0.9902500000000001
3	0.963499
4	0.9878225
5	

Class Score Fusion
(가중 평균값)



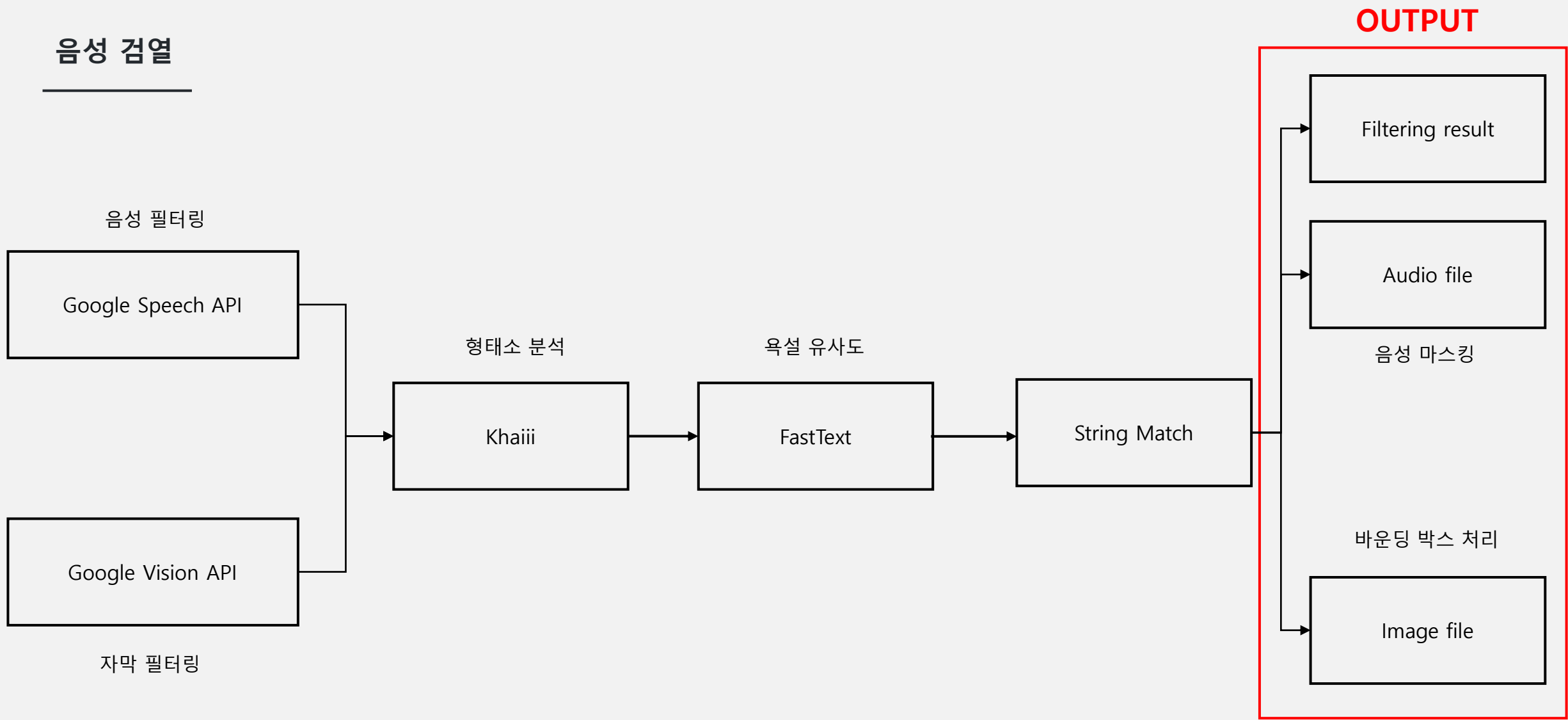
진행 현황 - 음성 및 자막 검열

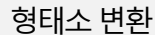
음성 및 자막 검열 진행도





음성 검열







욕설 검열 과정 (음성)

줄/VA 아요/EC
구/MAG 독/NNG
부탁/NNG 드리/VV 겠/EP 습니다/EC
안녕/NNG 하/XSA 세요/EC
신대리/NNG 이/VCP ㅂ니다/EC
다/MAG 들/XSN
반갑/VA 습니다/EC
오늘/NNG 은/JX
우리나라/NNG

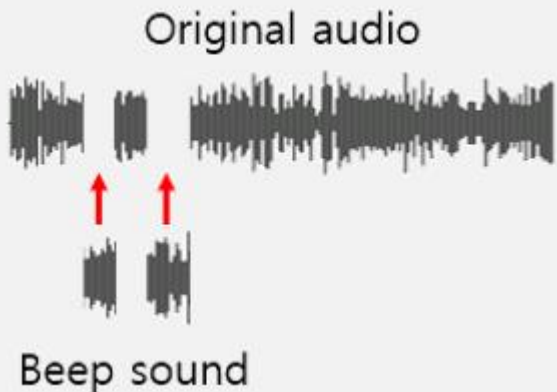
줄 0.07257053
아 요 0.4825888
구 0.018253189
독 0.26963842
부탁 0.3488102
드리 0.27959976
셋 0.04729164
습 니 다 0.40412754
새 끼 0.99999994
들 0.21808407
씨 발 새 끼 0.65116084
십 안 0.33072433

if (cosine similarity > 0.6):
String Match 진행

34
45
47
54
58
59
63
70
78
79
81
84

욕설 단어가 포함된 라인
(발화 시간 파일과 비교하여 시간 추출)

씨발새끼 / 0 : 21 / 0 : 22
새끼 / 0 : 24 / 0 : 24
씨발새끼야 / 0 : 27 / 0 : 28
개새끼야 / 0 : 28 / 0 : 29
좃 / 0 : 29 / 0 : 29





욕설 검열 과정 (자막)



자막 인식을 통해 텍스트와
단어의 좌표 저장

지쳤다! 너 너 너너 여기서 꺼져!!
어이 아줌마 잠시만! 예 쉼프! 세상에라고??
그림 중까 라고 해줄까??
블랙 자켓?? 염병 죄수복이 필요하겠지
하 씨발 이게 뭐야!
바비? 예 쉼프 지금이 역겹고 창피한 상황을 통제해 줄 누군가가 난 필요하다

0.jpg/지쳤다! 너 너 너너 여기서 꺼져!! +(88,219) (258,2
(126,238)/너+(133,219) (144,219) (144,238) (133,238)/너-
(193,220) (224,220) (224,239) (193,239)/꺼져+(226,220) +
2.jpg/어이 아줌마 잠시만! 예 쉼프! 세상에라고?? +(91,221
(156,239) (126,239)/잠시만+(160,220) (191,220) (191,239)
+(221,220) (240,220) (240,235) (221,235)/!+(243,224) (2
(182,255) (174,255)/??+(185,240) (208,240) (208,255) (1
5.jpg/그림 중까 라고 해줄까?? +(104,239) (229,239) (229,
(149,255)/라고+(163,243) (182,243) (182,255) (163,255)/i
(229,245) (229,254) (221,254)/



진행 현황 - 음성 및 자막 검열

욕설 검열 과정 (자막)

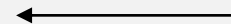
0.jpg/지쳤다! 너 너 너너 여기서 꺼져!! +(88,219) (258,219) (126,238)/너+(133,219) (144,219) (144,238) (133,238)/너- (193,220) (224,220) (224,239) (193,239)/꺼져+(226,220) (226,239) 2.jpg/어이 아줌마 잠시만! 예 쉼프! 세상에라고?? +(91,220) (156,239) (126,239)/잠시만+(160,220) (191,220) (191,239) +(221,220) (240,220) (240,235) (221,235)/!+(243,224) (243,239) (182,255) (174,255)/??+(185,240) (208,240) (208,255) (185,255) 5.jpg/그림 중간 라고 해줄까?? +(104,239) (229,239) (229,255) (149,255)/라고+(163,243) (182,243) (182,255) (163,255)/! (229,245) (229,254) (221,254)/



음성과 같은 방식으로
욕설 검열 진행

9.jpg/(141,243) (160,243) (160,255) (141,255) 17.jpg/(42,219) (64,219) (64,238) (42,238) 17.jpg/(18,240) (37,240) (37,255) (18,255) 17.jpg/(42,219) (64,219) (64,238) (42,238) 17.jpg/(18,240) (37,240) (37,255) (18,255)

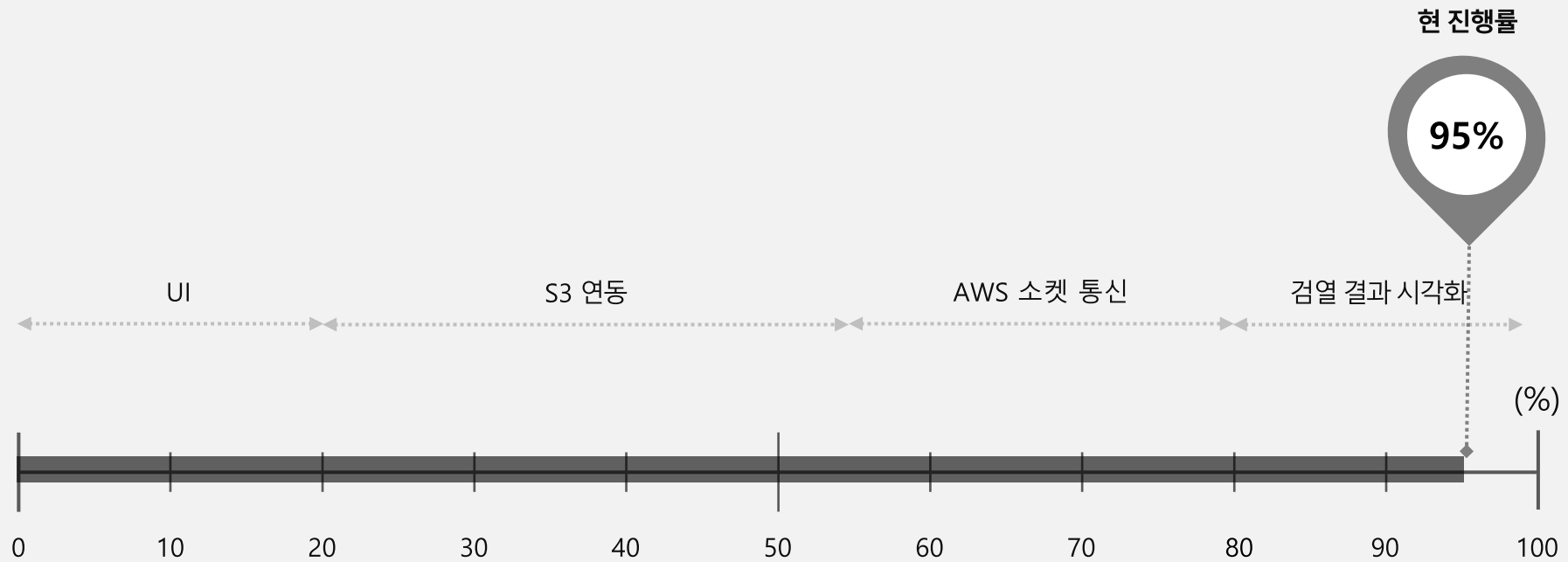
욕설 단어가 발견된 프레임과
해당 프레임 내의 좌표 추출

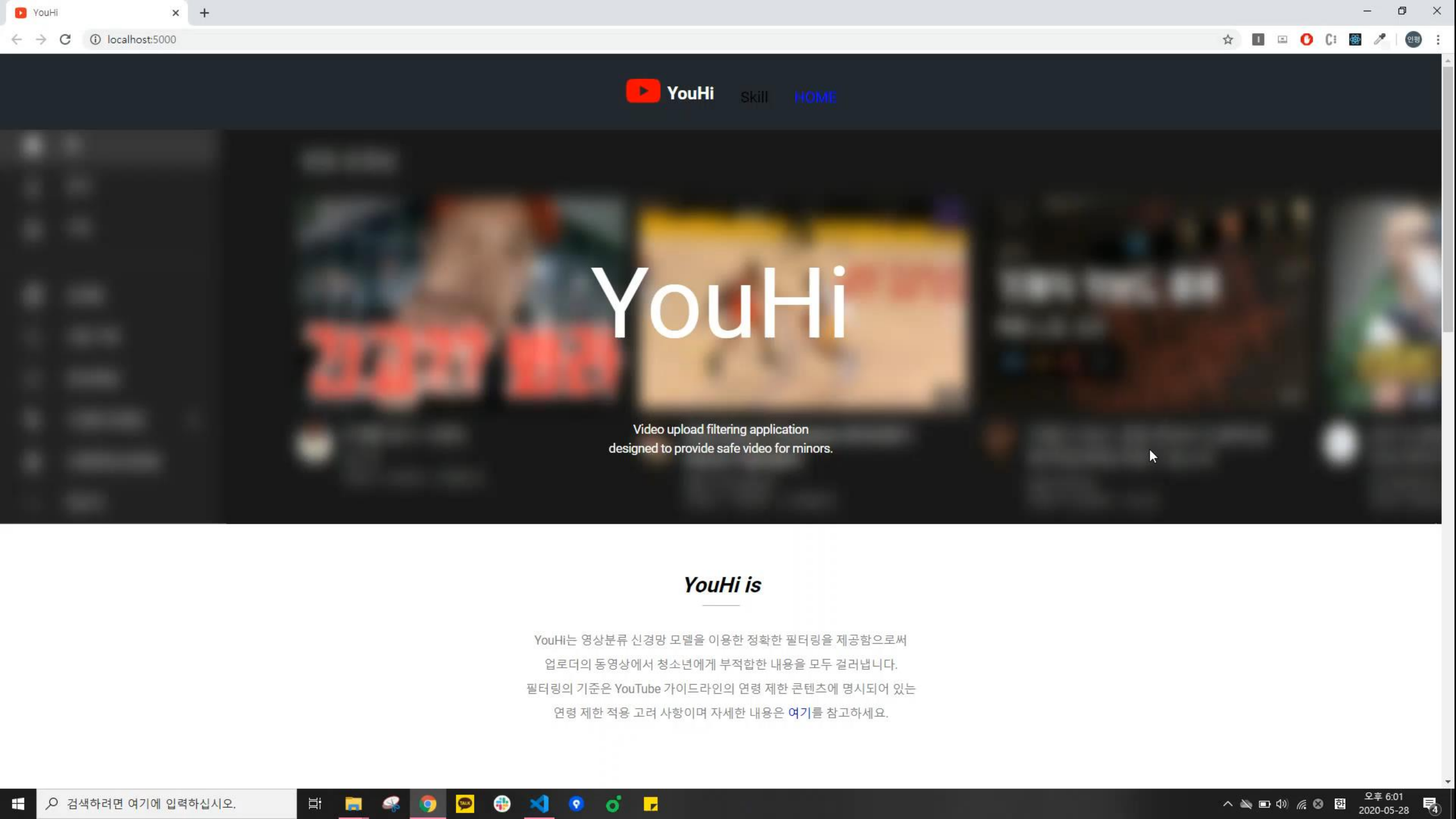


결과 출력



Front 및 Back-end 진행도





YouHi

Skill

HOME

YouHi

Video upload filtering application
designed to provide safe video for minors.

YouHi is

YouHi는 영상분류 신경망 모델을 이용한 정확한 필터링을 제공함으로써
업로더의 동영상에서 청소년에게 부적합한 내용을 모두 걸러냅니다.
필터링의 기준은 YouTube 가이드라인의 연령 제한 콘텐츠에 명시되어 있는
연령 제한 적용 고려 사항이며 자세한 내용은 [여기](#)를 참고하세요.



Front-end

- 로그인 / 이의 신청 기능
 - UI/UX 개선



Back-end

- 유지 보수

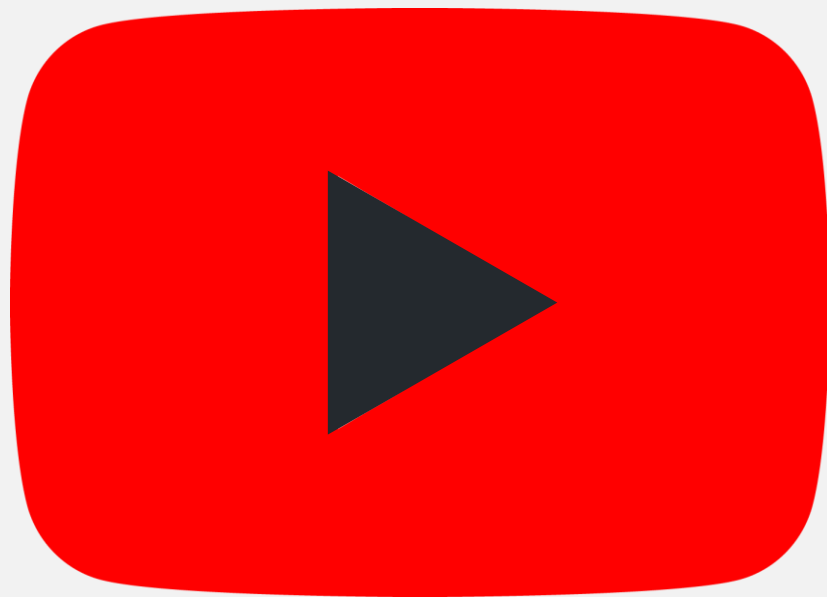


검열 기준 설정

- 모호한 YouTube 이용 정책
- 국가, 플랫폼별 상이한 심의기준

검열 시간

- 프레임 추출 및 구간 영상 추출
 - 총 4개의 딥러닝 모델



Thank you.

들어주셔서 감사합니다.