

2019 국립공원 빅데이터 분석 아이디어 공모전

생태계 교란 식물 실시간 탐지 및 국립공원 활성화 방안

- YOLO v2를 통한 이미지 데이터 학습 -

한석봉 팀
[홍한결, 석강, 방일한]

2019 혁신성장 청년인재 양성과정
실무 프로젝트 기반 빅데이터 전략 마에스트로 과정



CONTENTS

01 분석배경

- 분석 배경
- 분석 목적

02 데이터 수집

- 분석 계획
- 데이터 수집 및 전처리

03 모델 준비/평가

- 모델 소개
- 모델 구현
- 모델 평가

04 기대효과/제안

- 기대효과 & 활용방안
- 한계점 및 보안



CONTENTS

01

분석배경

- 분석 배경
- 분석 목적

02

데이터 수집

03

모델 준비/평가

04

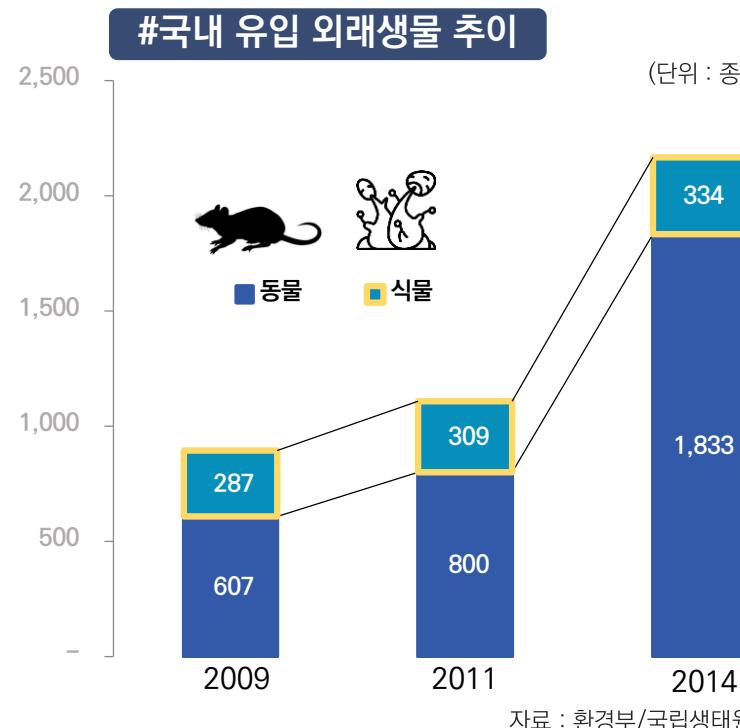
기대효과/제안

분석 배경 – 문제인식

생태계 교란 식물의 증가와 위해성

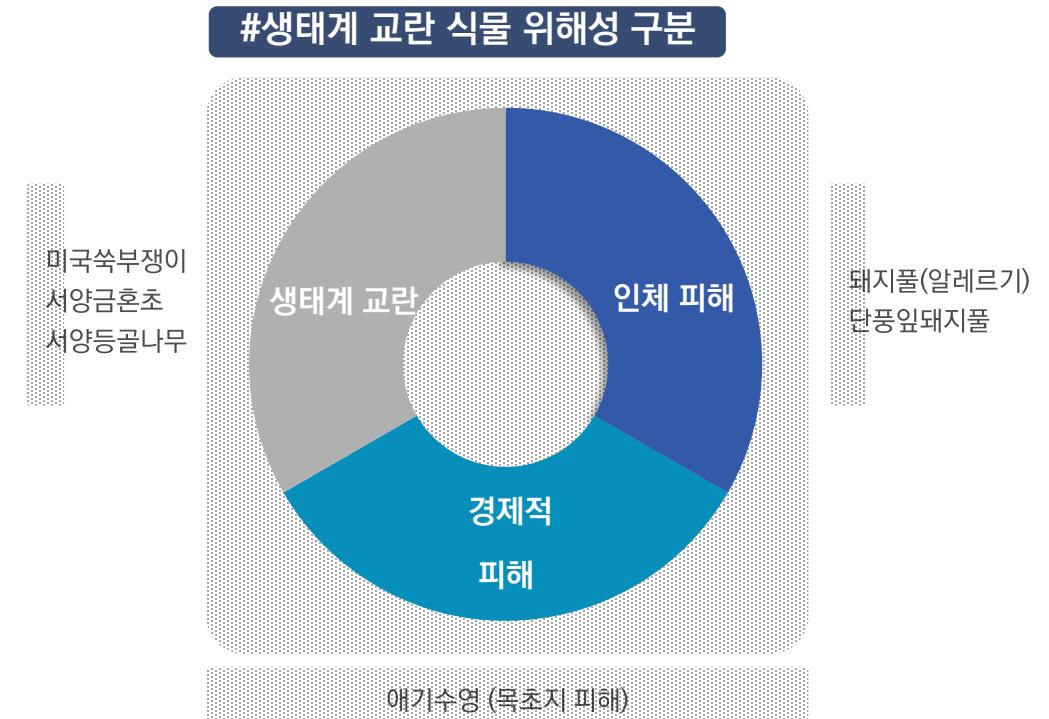
1. 생태계 교란 생물?

- 외래생물 중 생태계의 균형을 교란하거나 교란할 우려가 있는 생물



2. 생태계 교란 식물의 위해성

- 인체 피해 / 경제적 피해 / 생태계 교란



분석 배경 – 문제인식

국립공원의 생태교란 식물 관리
예방 / 제거 / 억제 방법으로 제거중

3. 생태계 교란 식물 지정 및 관리

- 국립공원 내 생태교란 야생식물은 13종이 분포하고 있으며, 전 국립공원에서 제거사업 실시 중

#생태교란 식물 지정 목록

구분	종명	학명	지정년도
식물 (14 종)	돼지풀	Ambrosia artemisiifolia	1999
	단풍잎 돼지풀	Ambrosia trifida	1999
	서양등골나물	Ageratina altissima	2002
	털물참새피	Paspalum distichum L. var. distichum	2002
	물참새피	Paspalum distichum	2002
	도깨비가지	Solanum carolinense	2002
	애기수영	Rumex acetosella	2009
	가시박	Sicyos angulatus	2009
	서양금흔초	Hypochaeris radicata	2009
	미국쑥부쟁이	Symphytum pilosum	2009
	양미역취	Solidago altissima	2009
	가시상추	Lactuca serriola	2012
	갯줄풀	Spartina alterniflora	2016
	영국겟끈풀	Spartina anglica	2016

자료 : 2017년 환경부

4. 국립공원의 외래식물 방제 기술

- 예방 / 제거 / 억제를 적절히 결합한 통합관리

#접근 방법에 따른 적용기술

접근 방법	적용 기술
예방	위해성평가, 조사 및 모니터링, 소독 등
제거	모든 방법을 이용한 완전 제거
억제	물리적 방법
	뿌리째 뽑기, 갈아엎기, 베기, 기계 이용 등
	화학적 방법
	제초제 살포
생물학적 방법	천적이용 : 곤충, 병원균
	가축방복
기타	인위적 불 놓기
	식생 복원

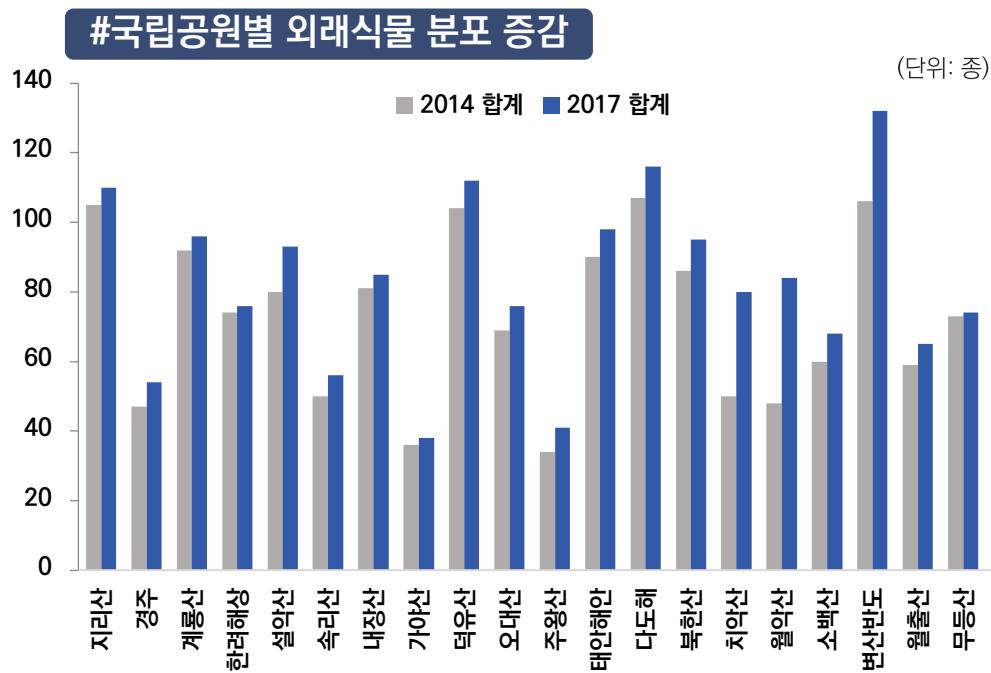
자료 : 2014년 특정 외래생물종 관리방안 연구(외래식물 분야)

분석 배경 - 문제인식

국립공원의 생태교란 식물 분포 및 제거 현황

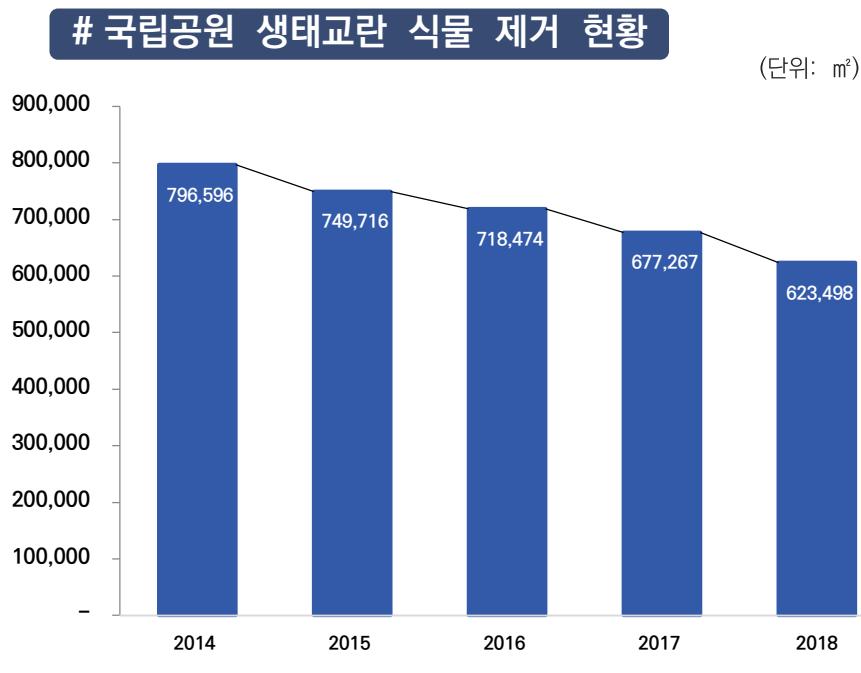
5. 국립공원별 외래식물(교란종 포함) 분포 현황

- 전체 국립공원 내 외래식물 종수가 증가
2014년 평균 72.5종 → 2017년 평균 81.5종



6. 국립공원 생태계 교란 식물 제거 현황

- 전체 국립공원 내 생태교란 식물 제거 면적 감소 추세



분석 목적 – 과제도출

생태교란 식물 관리의 문제점
많은 인력과 시간 소요

7. 생태계 교란 식물 제거의 어려움

- 많은 인력, 시간 소요



“국립공원관리공단에서는 교란생물로 인한 생태계 교란을 최소화하기 위해 지속적으로 분포조사 및 반복 제거작업을 시행하고 **많은 인력과 시간이 필요한 사항**”
국립공원자원 생태계교란 생물관리 – 국립공원공단



생태교란종 제거는…
‘사람 손으로 뽑는 게 가장 좋은데 예산이 너무 많이 들어가 현실적으로 한계…’
‘낙동강 하구 생태계교란종 전국 환경단체, 퇴치 대책 촉구’ – 연합뉴스

7-1. 현 외래식물 방제의 문제점

- 인력, 시간, 비용 소요, 토양교란

#접근 방법에 따른 적용기술의 문제점

접근 방법	적용 기술	문제점
예방	위해성평가, 조사 및 모니터링, 소독 등	-
제거	모든 방법을 이용한 완전 제거	인력/비용/시간 소요, 토양교란
억제	물리적 방법	뿌리째 뽑기, 갈아엎기, 베기, 기계 이용 등
	화학적 방법	제초제 살포
	생물학적 방법	천적이용 : 곤충, 병원균
	기타	가축방복
		인위적 불 놓기
		식생 복원

자료 : 2014년 특정 외래생물종 관리방안 연구(외래식물 분야)

분석 목적 – 과제도출

생태교란 식물 관리의 문제점
많은 인력과 시간 소요

7-2. 교란종 식물의 낮은 인지도와 구분의 모호함

- 관계자 / 전문가의 교육 및 설명 필요



개똥쑥과 돼지풀



“자원봉사자들이
쑥과 돼지풀 구분을 많이
헷갈려하고 있습니다.”



분석 목적 – 과제도출

교란종 제거에 효율성을 높일 수 있는 방안은??

8. 인력, 시간 투입 요소의 효율성 제고 필요

투입 시간

생태계 교란 식물 교육 및 안내

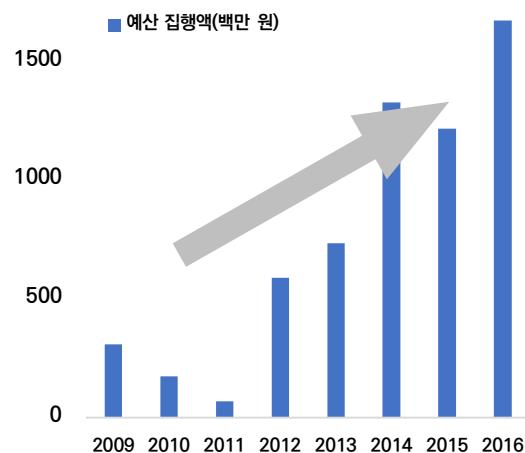
식물 유입, 확산 경로 추정 및 분석
지속적 관리 / 모니터링

투입 인력

생태계 교란종 전문가 / 관계자

체계적 교란종 관리를 위한 연구 및
제거 전문 위탁용역, 민/관/군 인력

강원도 생태계 교란식물 제거 예산 집행액 추이



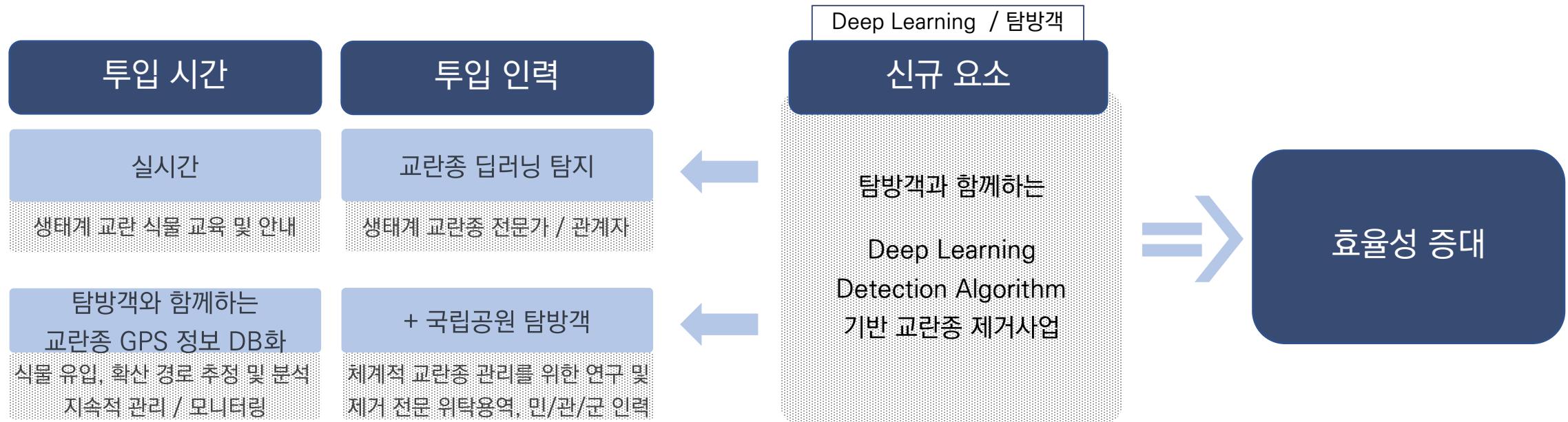
자료 : 2017 생태계 교란식물 관리 효율화 방안 연구, 강원연구원

효율성
제고안?

분석 목적 – 과제도출

생태계 교란종 실시간 탐지 알고리즘

9. Deep Learning Detection Algorithm 활용





CONTENTS

01
분석배경

02
데이터 수집

03
모델 준비/평가

04
기대효과/제안

- 분석 계획
- 데이터 수집 및 전처리

분석계획

높은 탐지 정확도, 교란·유사종 분류

1. 모델링 주요 포인트

- 이미지 : 선명 / 정확 / 충분한 수
- 알고리즘 : 높은 정확도로 탐지 / 분류

- 1) 교란종 인식 #국립공원 앱 내 야생화검색에서 돼지풀 인식 X(*별첨 참고)
- 2) 유사종 간 구분 - ex) 개똥쑥 / 돼지풀
- 3) 실시간 분석 - 관계자, 전문가 없이 교란종 인식이 가능!



교란종 관리에 필요한 시간/비용/인력 문제 해결, 효율성 증대

2. 업무 계획 및 분담표

- 업무 우선순위 및 팀원 별 강점을 바탕으로 분담

#업무 계획 및 분담표

진행 과정	자료 조사					
		데이터 수집 및 전처리				
		실시간 사물탐지 알고리즘 개발 및 구축				
					앱 서비스 제작	
팀원						
통합결	자료조사 / 업무계획 설정	자료조사	알고리즘 기획	앱 서비스 기획	앱 서비스 개발	앱 서비스 구현
석강		알고리즘 조사	알고리즘 기획	알고리즘 개발	알고리즘 구현	서비스 결과물 평가
방일한		자료조사	데이터 수집	데이터 수집 및 전처리	알고리즘 구현	서비스 결과물 평가

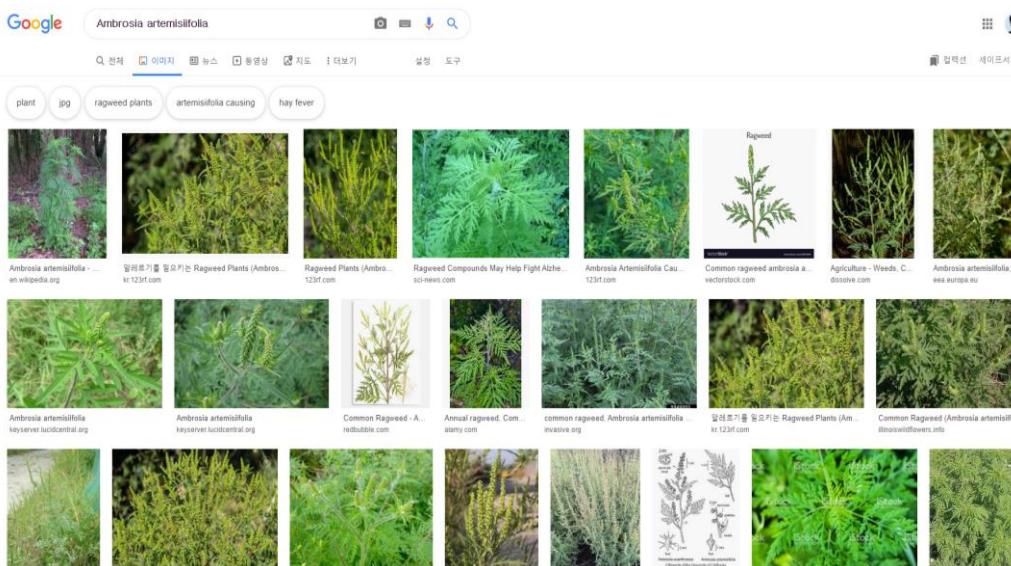
데이터 수집

식물 교란종에 대한 이미지 수집

3. 구글 이미지 크롤링

- Selenium 기반 이미지 크롤링 진행

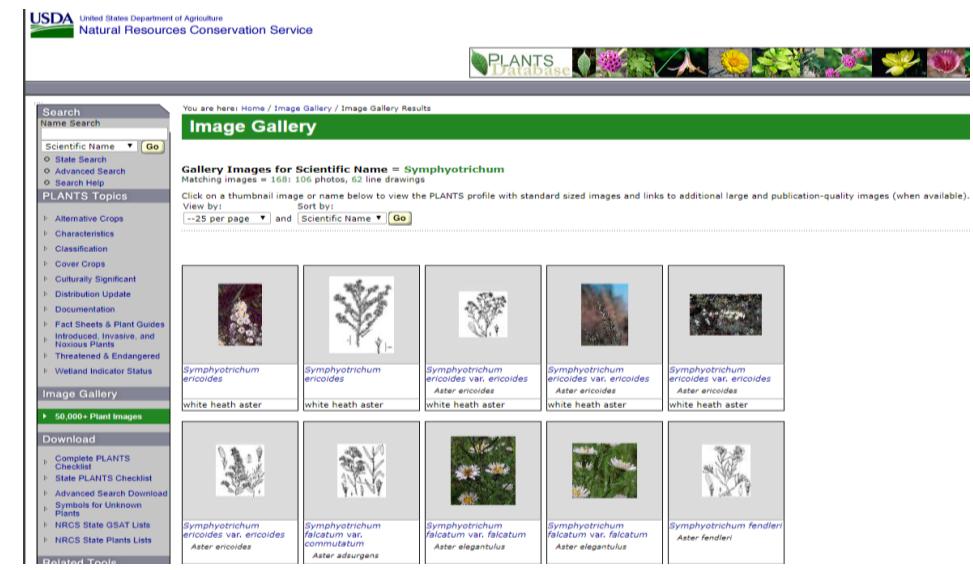
#구글 이미지 크롤링



4. 이미지 자체 수집

- 연구기관 / 정부기관 제공 이미지 수집(미국 농무부 USDA 등)

미국 USDA Plants Database



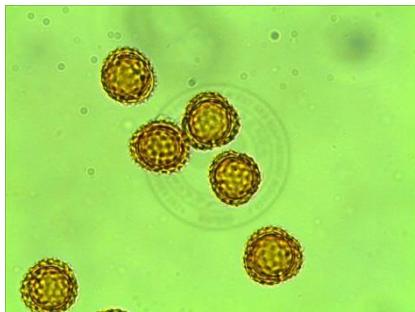
데이터 전처리

선명, 정확, 충분한 수의 이미지 데이터 확보

5. 이미지 삭제

- 불분명하고, 관련없는 이미지 데이터 삭제

#이미지 삭제



6. 이미지 데이터 추가 생성

- Generator (밝기 / 회전 / 상하좌우 반전 / 확대 등) 랜덤 적용



#Original Image



#Zoom



#Rotation



#Brightness



#Horizontal_Flip

데이터 수집 및 전처리 – 결과

유사종 간 구별과 실시간 인식을 위한 데이터셋 구축

7. 데이터 수집 및 전처리 결과

- 각 클래스 당 최소 200개 원본 데이터 / 800개 이상의 Generator 데이터 확보

#개똥쑥 VS 돼지풀

돼지풀 제거 시기 중 영양생장(4, 5, 6, 7월 – 집중제거 시기)의 이미지를 집중적으로 수집



개똥쑥 *Artemisia annua*
원본 – 207장
Generator – 828장

돼지풀 Common ragweed
원본 – 208장
Generator – 832장

#구절초 VS 미국쑥부쟁이



구절초 *latilobum*
원본 – 286장
Generator – 858장



미국쑥부쟁이 frost aster
원본 – 444장
Generator – 888장



CONTENTS

01
분석배경

02
데이터 수집

03
모델 준비/평가

04
기대효과/제안

- 모델 소개
- 모델 구현
- 모델 평가

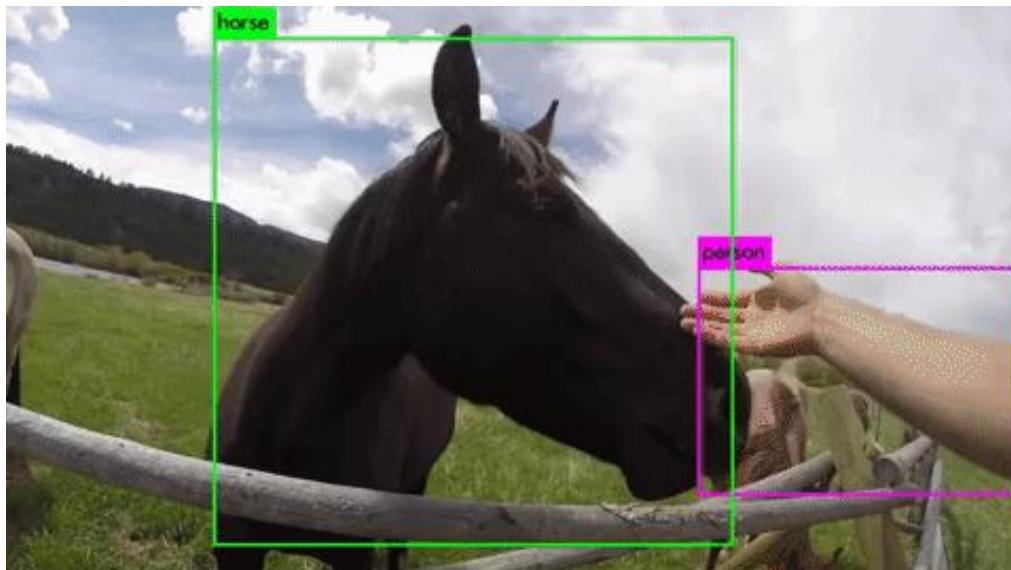
모델 소개 - YOLO v2 설명

YOLO v2 사용 - 높은 탐지 정확도 / 빠른 프레임 속도 / 실시간 이미지 분류

1. YOLO(You Only Look Once)

- Deep Learning Detection Algorithm

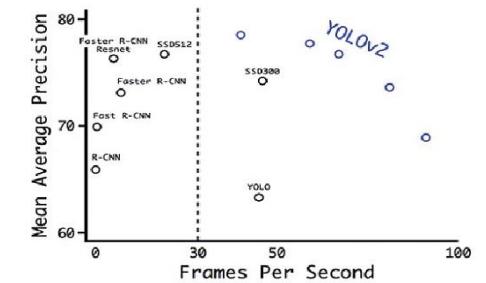
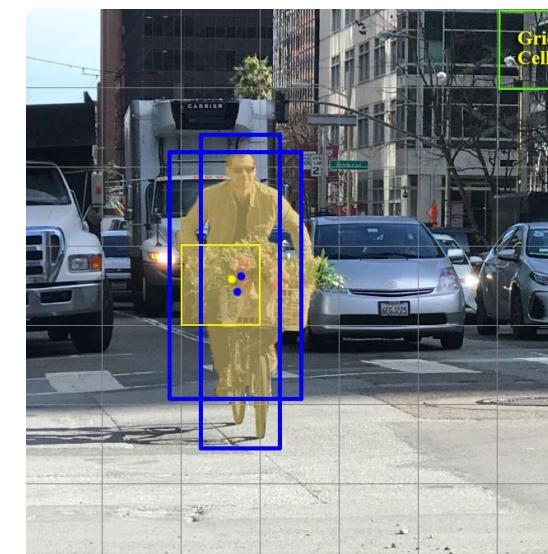
#YOLO 소개.gif



2. YOLO v2의 특징

- 빠른 프레임 속도와 높은 정확도, 객체 탐지 및 이미지 분류

#Deep Learning Detection Algorithm 간 성능 비교



Detection Frameworks	Train	mAP	FPS
Fast R-CNN [5]	2007+2012	70.0	0.5
Faster R-CNN VGG-16[15]	2007+2012	73.2	7
Faster R-CNN ResNet[6]	2007+2012	76.4	5
YOLO [14]	2007+2012	63.4	45
SSD300 [11]	2007+2012	74.3	46
SSD500 [11]	2007+2012	76.8	19
YOLOv2 288 × 288	2007+2012	69.0	91
YOLOv2 352 × 352	2007+2012	73.7	81
YOLOv2 416 × 416	2007+2012	76.8	67
YOLOv2 480 × 480	2007+2012	77.8	59
YOLOv2 544 × 544	2007+2012	78.6	40

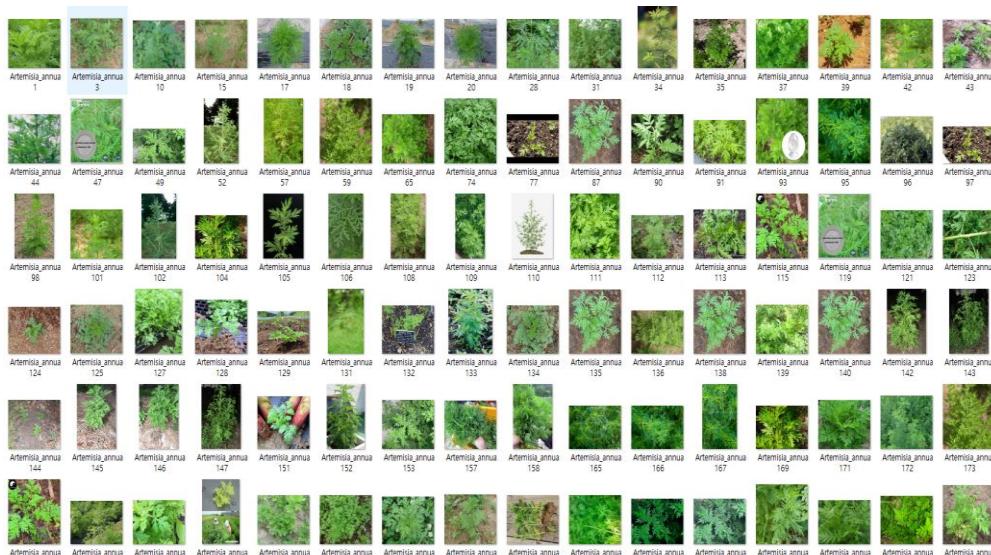
모델 구현 - 학습과정

4개 클래스 Image Boundary-Box labeling 진행

3. 학습 데이터 준비

- 생태교란종 이미지 데이터 총 1,145장

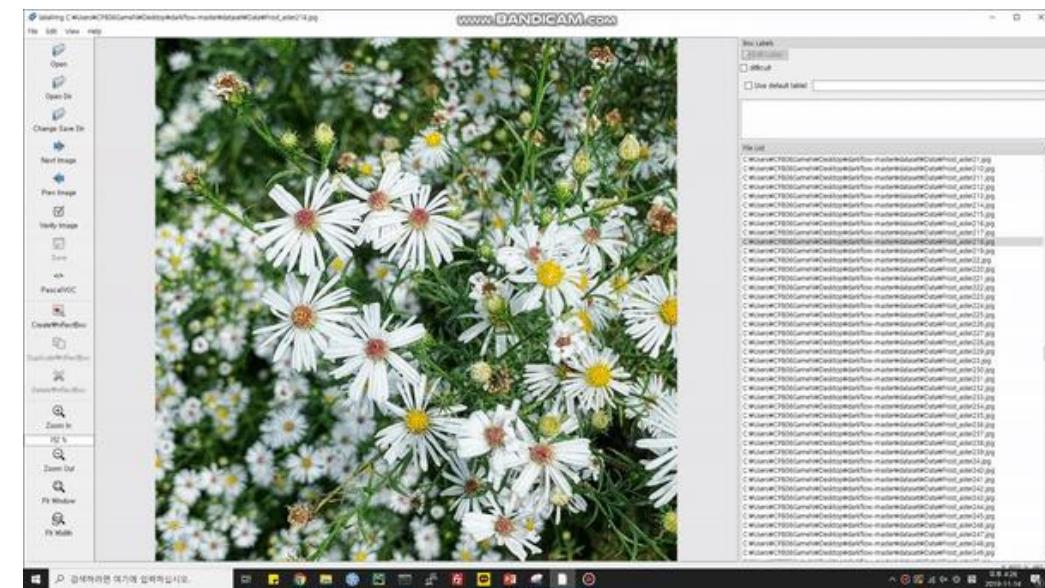
#YOLO v2 학습 데이터셋



4. 이미지 라벨링 과정

- Image boundary-box labeling

#이미지 라벨링.gif



모델 구현 - 학습과정

52시간, 총 2일 4시간 학습 시행

5. YOLO v2 학습

- 45,000번 시행 / 수행시간 52시간

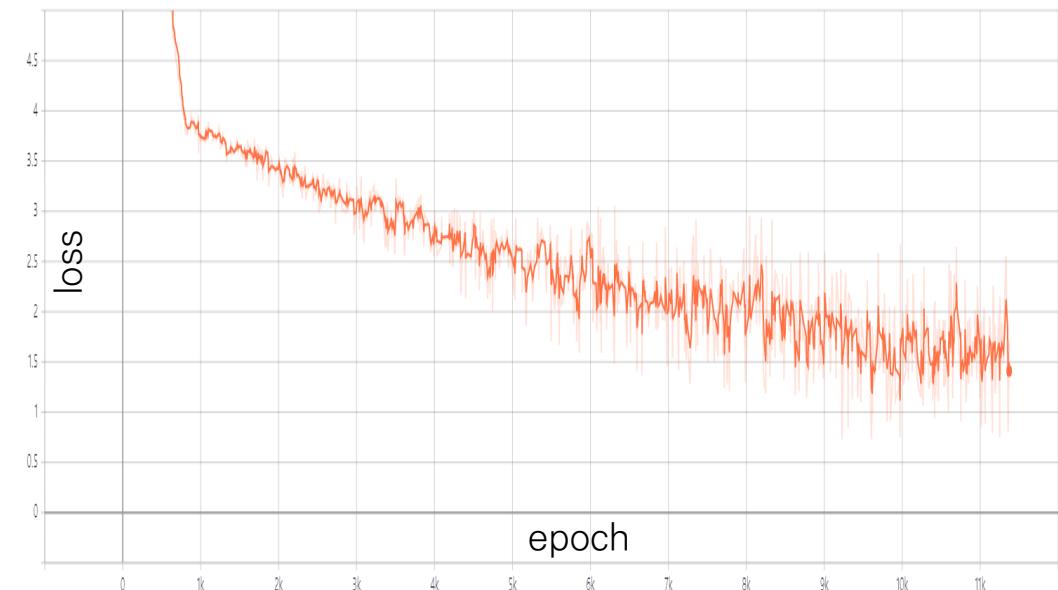
#YOLO v2 학습 진행

```
step 1847 - loss 6.014066696166992 - moving ave loss 5.341483371078816
step 1848 - loss 5.13972806930542 - moving ave loss 5.3213078409014765
step 1849 - loss 4.2864837646484375 - moving ave loss 5.217825433276173
step 1850 - loss 6.381662368774414 - moving ave loss 5.334209126825996
Checkpoint at step 1850
step 1851 - loss 5.543975830078125 - moving ave loss 5.35518579715121
step 1852 - loss 4.70591926574707 - moving ave loss 5.290259144010796
step 1853 - loss 4.736455917358398 - moving ave loss 5.234878821345557
step 1854 - loss 3.6587491095461426 - moving ave loss 5.077265849565616
step 1855 - loss 5.46517038345369 - moving ave loss 5.116056302954392
step 1856 - loss 6.018374443054199 - moving ave loss 5.206288116964372
step 1857 - loss 5.407800197601318 - moving ave loss 5.226439325028067
step 1858 - loss 6.2400312423706055 - moving ave loss 5.3277985167623205
step 1859 - loss 4.549107551574707 - moving ave loss 5.249929420243559
step 1860 - loss 5.049672603607178 - moving ave loss 5.229903738579921
Checkpoint at step 1860
step 1861 - loss 6.092799186706543 - moving ave loss 5.316193283392583
step 1862 - loss 6.020009517669678 - moving ave loss 5.386574906820293
step 1863 - loss 4.861415863037109 - moving ave loss 5.334059002441975
step 1864 - loss 4.878724575042725 - moving ave loss 5.28852555970205
step 1865 - loss 5.662031173706055 - moving ave loss 5.325876121102451
step 1866 - loss 4.612056255340576 - moving ave loss 5.254494134526264
step 1867 - loss 6.0909833908081055 - moving ave loss 5.338143060154449
step 1868 - loss 6.324238300323486 - moving ave loss 5.436752584171353
step 1869 - loss 5.761890888214111 - moving ave loss 5.469266414575629
```

6. YOLO v2 loss Visualization

- 최종 loss 0.93451

#YOLO v2 loss 추이



모델 평가 - 학습결과 / 평가

7. 교란종 인식

- 돼지풀(Common ragweed)
- 정확도 0.51 ~ 0.54

05. 뉴스영상_돼지풀_dectection.mp4

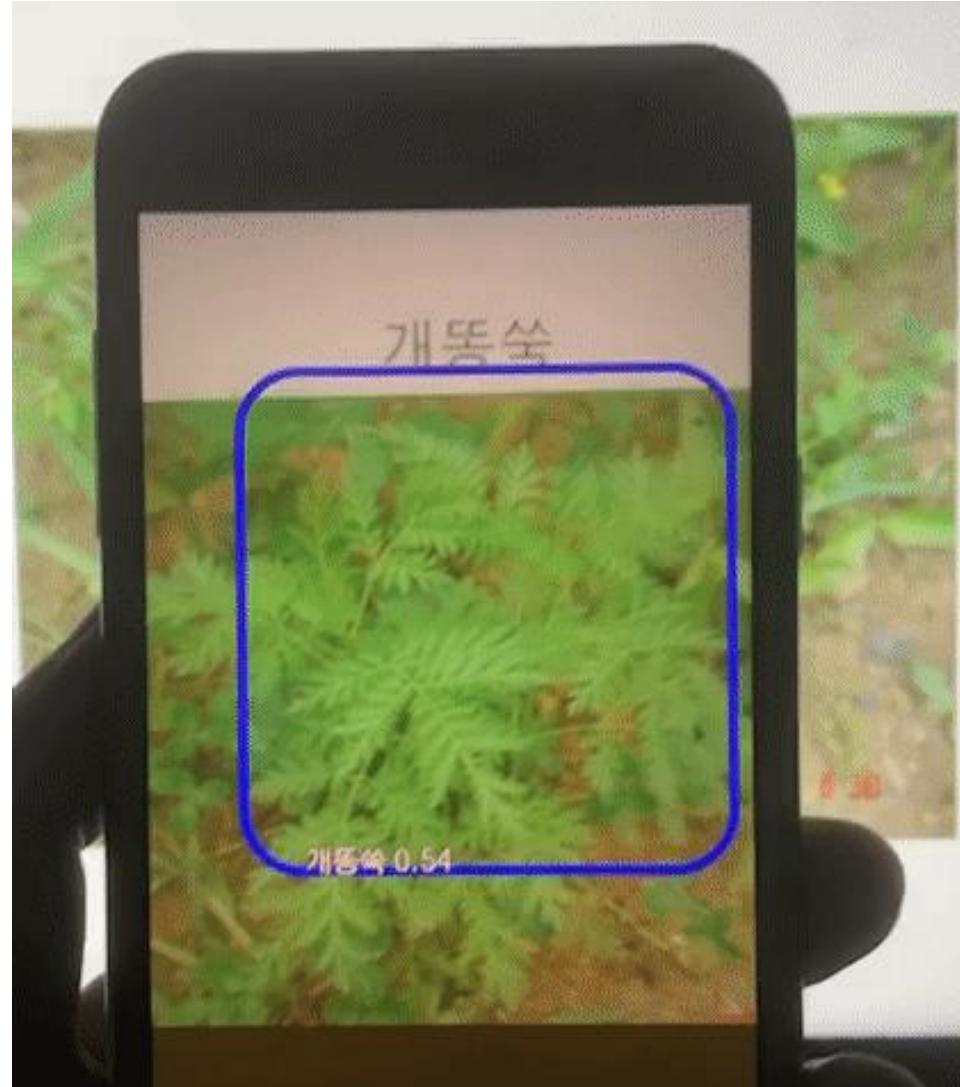


모델 평가 - 학습결과 / 평가

8. 교란종 유사종 간 구분

돼지풀 vs 개똥쑥 구분

- 돼지풀 VS 개똥쑥
- 정확도 0.48 ~ 0.71





CONTENTS

01
분석배경

02
분석과정

03
모델 준비/평가

04
기대효과/제안

- 솔루션 기획
- 솔루션 구체화
- 한계점 및 보완

솔루션 기획 – 니즈분석

교란종 식물에 대한 탐방객의 인식 제고, 지속적 모니터링 필요

1. 교란종에 대한 탐방객 인식 제고 필요

- 탐방객의 인식 제고 및 외래(교란종)식물 정보 공유의 필요성

교란종 식물 관리방안 및 제안_1

4 초기 확산 방지 및 체계적 관리 필요

- 전 공원의 외래식물 분포 현황에 대한 데이터가 부족한 실정이며, 외래식물은 새로운 장소에 정착하여 확산되면 제거 및 관리하는데 인력, 시간 및 비용이 많이 소요되므로 초기에 확산을 방지하는 것이 가장 효과적임
- 국립공원에서 외래식물을 확산을 방지하기 위해 에어컴프레셔, 스틸그레이팅, 매트를 설치하여 운영하고 있음. 하지만, 무엇보다 탐방객의 인식을 제고할 필요성이 있으며, 외래식물에 대한 정보를 공유하고 참여할 수 있는 방안이 필요함
- 매년 우선 공원을 선정하여 체계적으로 분포 현황 및 제거방법, 생물학적 방제방법 등을 고려하여 공원별 대체 식재 종을 선정하고 체계적이고 지속 가능한 외래식물 관리시스템을 구축하여 방제를 실시해야 함

2. 발생지의 지속적인 모니터링

- 외래(교란종)식물 발생지의 주기적 조사 및 모니터링 필요

교란종 식물 관리방안 및 제안_2

5 지속적인 모니터링 필요

- 외래식물은 군락형태의 생육상을 보이므로 한번 발생하여 확산되면 제거하기 어려우며, 기존에 외래식물 발생지는 주기적인 조사 및 모니터링이 필요함

솔루션 기획 - 활용 아이디어

실시간 교란종 탐지 기술을 활용한
인식 제고 / 지속적인 모니터링

3. 교란종 실시간 탐지 기술로 인식 제고

- 교란종 탐지 기술 홍보 / 제거 사업 시 기술 활용

교란종 실시간 탐지

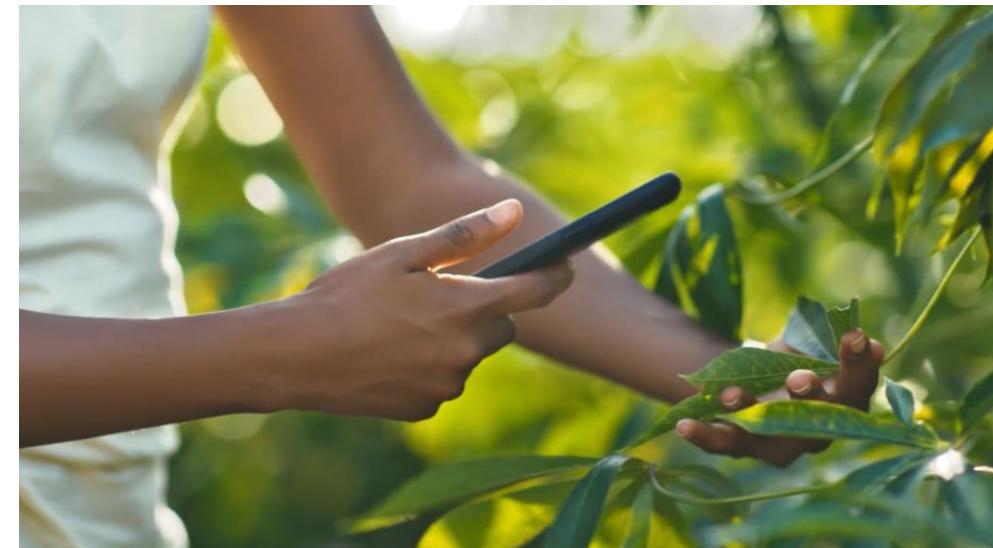


자료 : TensorFlow, an ML platform for solving impactful and challenging problems

4. 탐방객과 함께하는 교란종 GPS 정보 공유

- 외래식물 발생지의 GPS 정보의 DB화 및 실시간 모니터링 체계 구축

교란종 사진 촬영 후, GPS 데이터 공유



자료 : TensorFlow, an ML platform for solving impactful and challenging problems

솔루션 구체화 – 타겟 설정

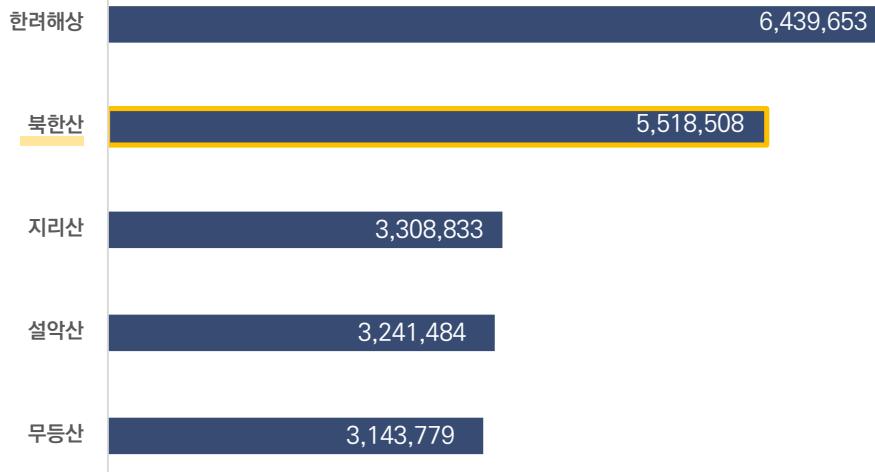
‘북한산’ 탐방객 / 심각성

5. 2018년 국립공원별 탐방객 현황

- 해상국립공원 제외 시, 육상국립공원 내 북한산이 탐방객 수 1위

2018년 국립공원별 탐방객 현황 TOP5

(단위: 명)



자료 : 2019년 국립공원 기본 통계

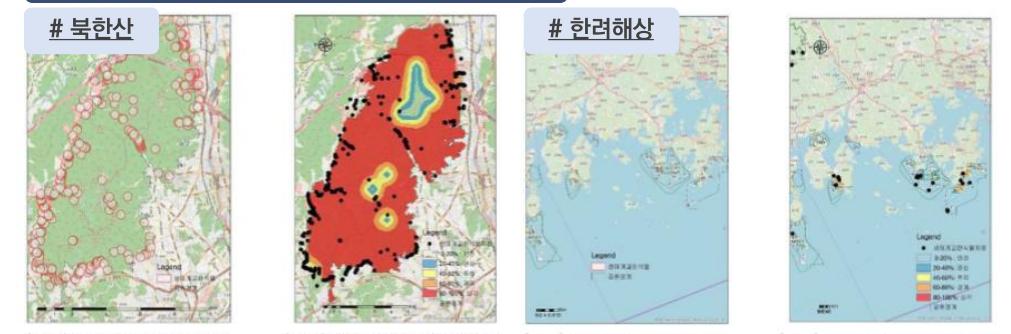
6. 국립공원별 교란식물 분포 현황 및 확산 위험 예측도

국립공원 생태계교란식물 분포 현황 TOP5

국립공원	교란식물 종수	면적(m^2)	분포지점
북한산	6	123,622	1,574
지리산	2	12,881	164
설악산	4	6,205	79
무등산	4	4,555	58
한려해상	2	3,848	49

2017년까지 국립공원에서 조사된 식물

교란식물 확산 예측도 및 평가표



자료 : 2017년 국립공원 외래식물 관리방안 연구

솔루션 구체화 – 캠페인 아이디어 제시

환경 지키미 캠페인
여드름(Yo! drop)

7. 여드름(Yo! Drop)?

- 국립공원 탐방객 참여형 교란종 관리 캠페인
- 우리 자연에 대한 주인의식을 가지고 참여하는 환경 캠페인

“우리 모두가 함께 지켜 나가는 국립공원”



8. 여드름(Yo! Drop) 캠페인 주제

“사람들의 피부 속 여드름처럼
우리나라의 생태계를 해치는 교란종들을 제거하자”



“생태계 교란종 = 여드름”

솔루션 구체화 - 캠페인 사업 구체화

여드름(Yo! drop)

그린포인트 연계 시민참여형 교란종 제거 사업 실시

9. 교란종 제거 사업

- 실시간 생태계 교란종 인식 App 활용
- 시민참여형 교란종 제거 사업 실시

시민참여형 교란종 제거 사업



10. 그린포인트 연계

- 그린포인트란? - 국립공원 내 버려진 쓰레기 수거 및 자기 쓰레기를 되가져가는 경우
포인트를 제공받고, 이를 공원시설에 이용, 상품으로 교환하는 문화활동
- **수매사업의 효율성을 토대로 그린포인트와 연계한 교란종 제거**
[명사] 수매 – 거두어 사들임. 또는 그런 일.

사업 방식별 생태계 교란식물 제거면적 및 효율 비교

구 분	직접 인력고용	전문업체 위탁용역	수매사업
사업단가(원/m ²)	130	116	180
예 산 (원)	계(원)	1,750,000,000	1,750,000,000
	인건비	1,141,276,779	1,108,888,674
	기 타	608,723,221	18뿌리/m ² 10원/뿌리
예상제거면적(m ²)	8,789,795	9,559,385	9,722,222
직접 인력고용 대비 효율(배)	1.00	1.09	1.11

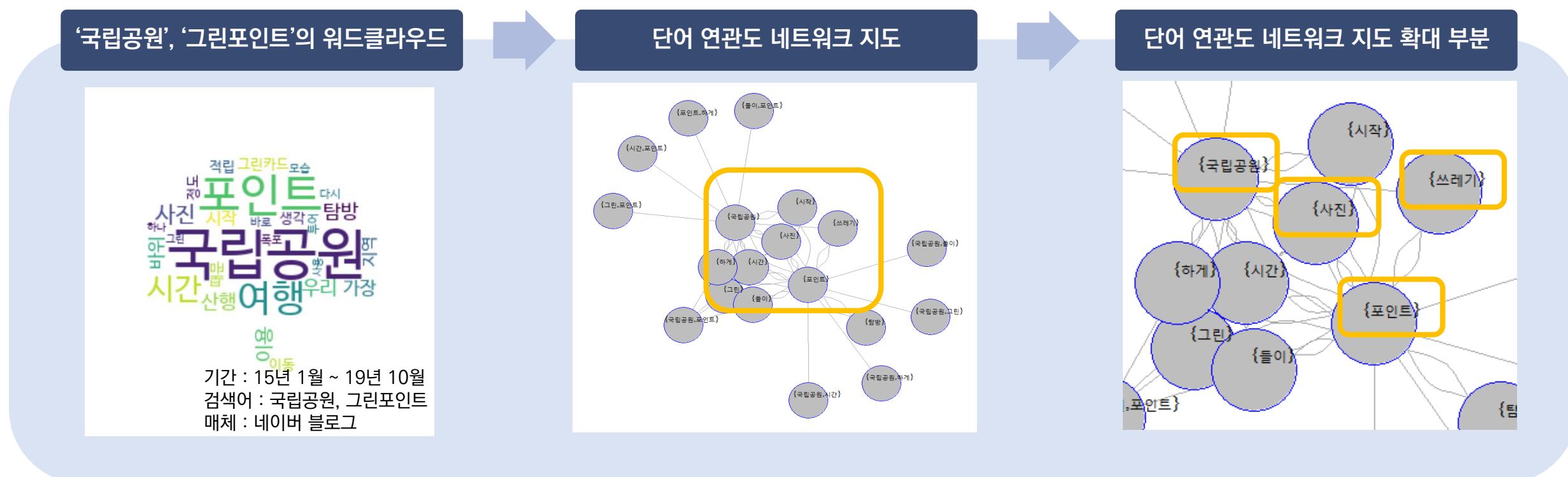
자료 : 2017 생태계 교란식물 관리 효율화 방안 연구, 강원연구원

인위적인 재배 및 판매 등 모럴해저드(Moral hazard) 발생을
고려하지 않을 수 없으므로 제도적인 측면 강구 필요

솔루션 구체화 - 캠페인 사업 구체화

여드름(Yo! drop)
국립공원, 그린포인트, 사진, 쓰레기

11. 국립공원/그린포인트 연관어 도출



솔루션 구체화 - 캠페인 사업 구체화

여드름(Yo! drop)

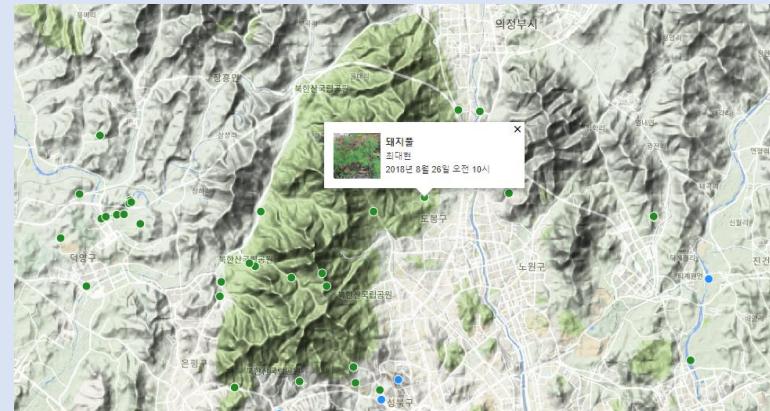
그린포인트 연계 국립공원 탐방객의 교란종 GPS 정보 공유

12. 교란종 사진을 통한 GPS 시스템 구축 + 그린포인트 제도 연계

교란종 인식 시 사진촬영

GPS 정보 획득 / 구축 및 지도화

그린포인트 적립



네이처링(온라인 기반 자연활동 공유 플랫폼) 돼지풀 분포 지도 캡처



자료 : BGF

한계점 및 보완

1. 계절별 다양한 식물 모습 반영 부족 ex) 돼지풀 집중제거 시기(영양생장) 4,5,6,7월
 - > 캠페인 진행 시, 시민들이 촬영한 사진을 추가하여 모델 학습을 진행하면 더 정확한 인식률 확보
2. 5개 국립공원에서 공통적으로 출현하는 돼지풀을 최우선으로 분류, 이에 다른 교란식물 및 유사종 데이터 X
 - > 추가하지 않은 다른 교란식물 및 유사종 데이터 수집 및 학습 진행 필요
3. 모델 학습에 최소 48시간 소요 -> 모델링 수정 시 많은 시간 필요 -> 프로젝트 기간 내 탐지 및 분류 정확도 향상에 한계
 - > 충분한 시간 내 모델링 수정 진행 / 유사종 간 분류 안될 시, 추가 정보 제시(팝업 메시지) EX) 텔

참고목록 / 사용프로그램

참고목록 :

2019년 국립공원기본통계

2017년 국립공원 외래식물 관리방안 연구

2017년 생태계 교란식물 관리 효율화 방안 연구, 강원연구원

2016년 서울시 외래식물 분포특성과 관리방안 Ⅲ, 서울연구원

2009년 생태계교란 야생동식물 자료집

Real-time Object Detection with YOLO, YOLOv2 and now YOLOv3, @jonathan_hui

Deep Learning(YOLO)기반의 Smart Scarecrow, 개발자 이상현 블로그

Android에서 내가 학습한 YOLO 모델을 이용해 Object Detection 구현하기 <https://junyoung-jamong.github.io/>

사용프로그램 및 패키지 :



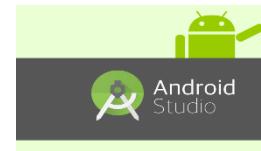
- Python
matplotlib, numpy, pandas,
pytagcloud, Selenium 패키지



- R
Rcpp, KoNLP, arules,
igraph 패키지

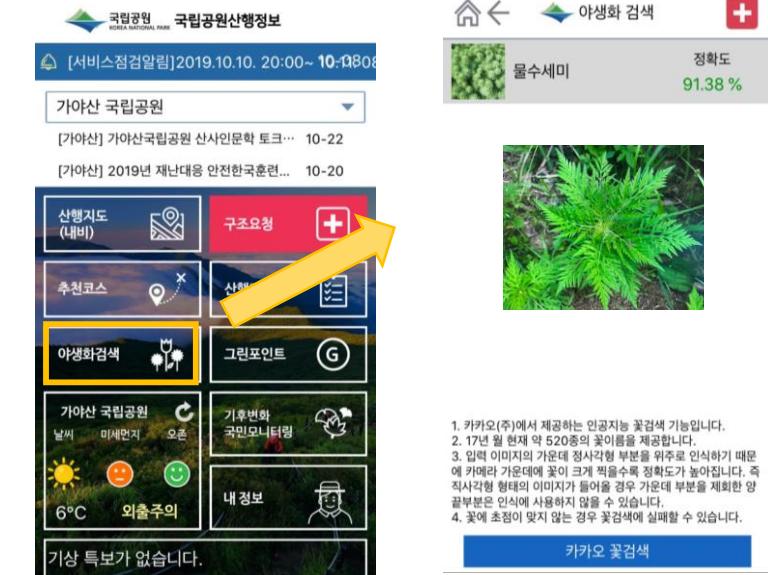


- TensorFlow
Darkflow library
YOLO model Deep learning



- Android Studio
Launch Application

별첨 - 국립공원 앱 내 야생화 검색



- 카카오(주)에서 제공하는 인공지능 꽃검색 기능입니다.
- 17년 월 현재 약 520종의 꽃이들을 제공합니다.
- 입력 이미지의 가운데 정사각형 부분을 위주로 인식하기 때문에 카메라 가운데에 꽃이 크게 찍을수록 정확도가 높아집니다. 즉 정사각형 형태의 이미지가 들어올 경우 가운데 부분을 제외한 양 끝부분은 인식에 사용하지 않을 수 있습니다.
- 꽃에 초점이 맞지 않는 경우 꽃검색에 실패할 수 있습니다.

카카오 꽃검색

Q & A

THANK YOU

