

딥러닝을 활용한 독성 약재 감별 정보 시스템

독약**한방**

A red explosion icon with black outlines and radiating lines, positioned above the right side of the main title.

2019. 8. 14
발표자 조재성



목차 INDEX

01 약재 이미지와 정보 데이터
데이터 수집

03 모델 선정 및 관리
모델
선정 및 관리

00 프로젝트 소개
프로젝트 배경

02 데이터 전처리와 설명가능 모델을 위하여
데이터 가공

04 UI에 따른 기대효과
UI 및
기대 효과

독약한방

Deep Learning을 활용한
독성 약재 감별 시스템

일반인용

전문가용

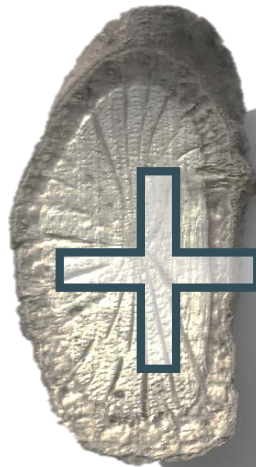
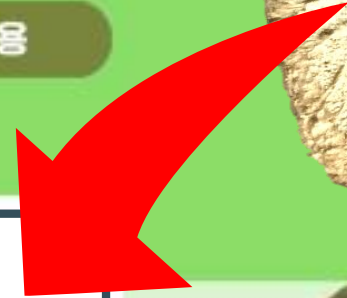


독약한방

Deep Learning을 활용한
독성 약재 감별 시스템

일반인용

전문가용



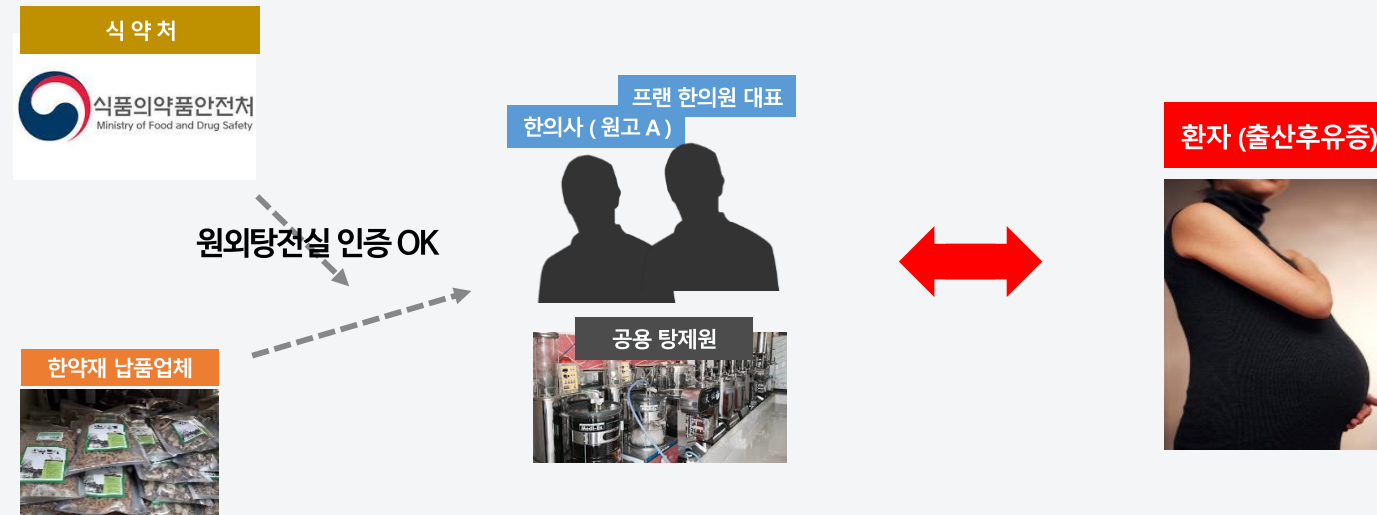
매년 발생하는 산행인들의 독약 흡입으로 인한 사망 (매년 발생하는 와중, 불과 2개월 전, 또...)

독성(毒性) 주의 한약재 복용 70대남성 사망

[2019년 06월 06일 16시 34분]

[데일리메디 박정연 기자] 대한한 의사협회(회장 최혁용)는 최근 70대 남성이 '초오(草烏)'를 넣어 끓인 국을 먹고 목숨을 잃은 사건과 관련, 독성(毒性)이 강한 한약재 복용시 한의사 진단과 처방에 따를 것을 당부.

심심하면 발생하는 한의원과 환자와의 법정 공방 (?? 식약처 인증 받았는데 ??)



❖ 불필요한 암기... 외워도 잊어 버릴텐데... 딥러닝으로 **독약을 한방**에 감별하고 정보를 제공하면 어떨까?



Akebiae Caulis



Aristolochiae
Manshuriensis Caulis



Sinomenium
acutum

6장. 이수삼습약-제 1절 이수퇴종약

茯苓	Hoelen	잔나비결상과	진균	복령	Poria cocos	菌核
猪苓(1)	Polyporus	잔나비결상과	진균	猪苓	Polyporus umbellatus	菌核
泽泻(2)	Alismatis Rhizoma	택사과	多年生 沼澤植物	질경이택사	Alisma plantago-aquatica	塊莖
薏苡仁	Coicis Semen	벼과	一年生 또는 多年生 草本	율무	Coix lachyma-jobi	種仁
冬瓜皮	Benincasae Pericarpium	박과	一年生 草本	동아호박	Benincasa hispida	果皮
赤小豆	Phaseoli Angularis Semen	콩과	一年生 草本	팥	Phaseolus angularis	種子
玉米鬚	Maydis Stigma	벼과	一年生 草本	옥수수	Zea mays	花柱

6장. 이수삼습약-제 2절 이수룡림약

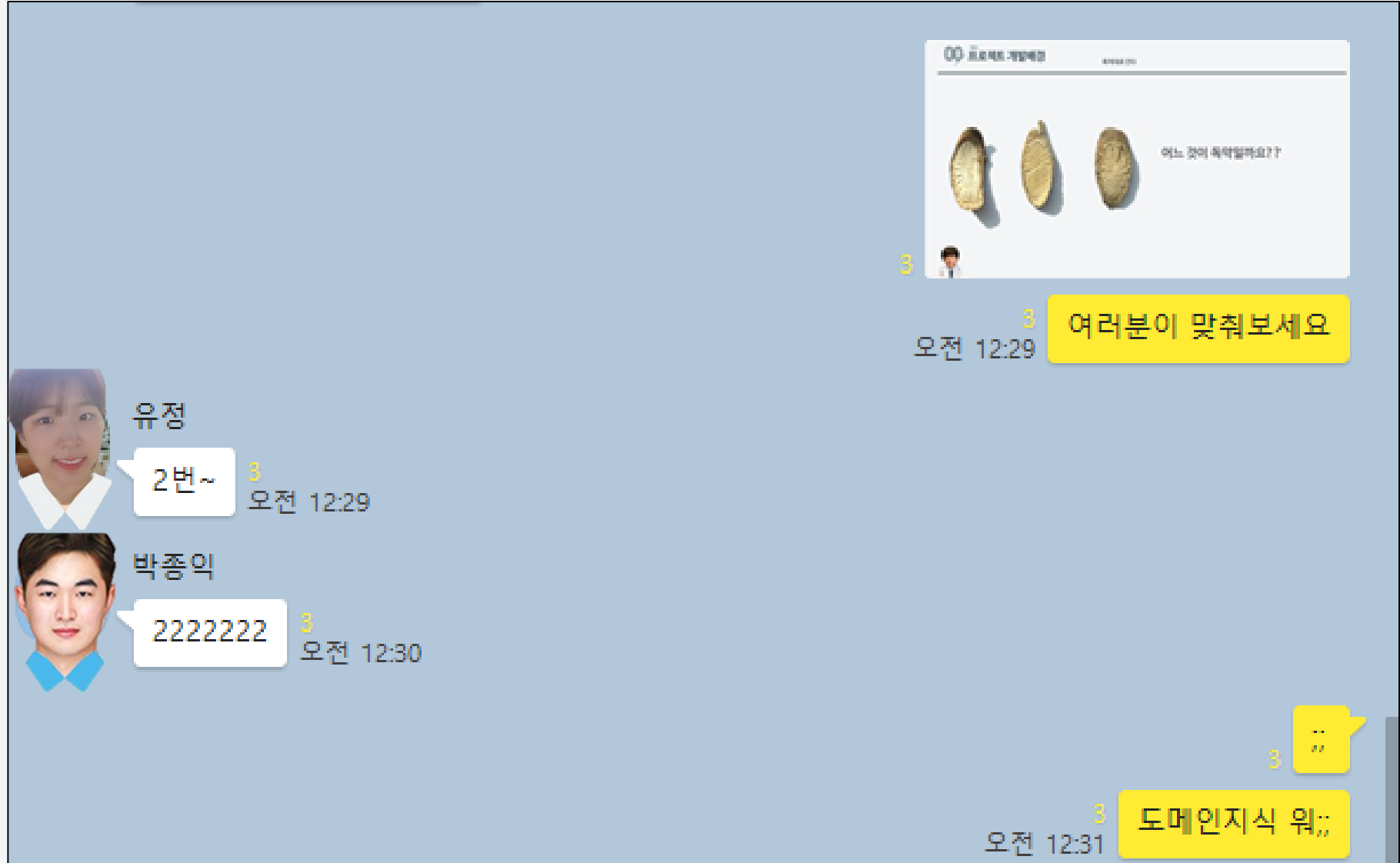
車前子(1)	Plantaginis Semen	질경이과	多年生 草本	질경이	Plantago asiatica	種子
木通(2)	Akebiae Caulis	프롤딩굴과	낙엽성 관목	프롤딩굴 소목통 삼엽목통	Akebia quinata Clematis armandii Akebia trifolia	□ 줄기
滑石(1)	Talcum	구산염류 광물	-	-	Mg ₃ (Si ₄ O ₁₀)(OH) ₂	
通草(2)	Tetrapanax Medulla	두릅나무과	관목	통말목(通脫木)	Tetrapanax papyriferus	莖髓
海金沙	Lygodii Spora	실고사리과	-	실고사리	Lygodium japonicum	孢子
石韋	Pyrrosiae Folium	고란초과	多年生 草本	석취 세활석취	Pyrrosia lingua P. tricuspidata	葉
葶藶	Toriko Rhizoma	마과	-	도프도마	Dioscorea tokoro	根莖
地膚子	Kochiae Fructus	명아주과	一年生 草本	명아주	Kochia scoparia	果實
篇蓄	Polygoni Avicularis Herba	마디풀과	一年生 草本	마디풀	Polygonum aviculare	全草
瞿麥	Dianthi Herba	식취과	多年生 草本	솔피명이꽃 파명이꽃	Dianthus superbus D. chinensis	全草
冬葵子	Malvae Semen	아욱과	一年生 草本	아욱	Malva verticillata	種子
燈心草	Junci Medulla	꽃참과	多年生 草本	꽃참	Juncus effusus	根髓
三白草(1)	Saururi Herba seu Rhizoma	삼백초과	多年生 草本	삼백초	Saururus chinensis	根, 全草
茵陳蒿(1)	Artemisiae Capillaris Herba	국화과	多年生 草本	사질썩(白蒿) 비썩(茵蒿) 더위저기	Artemisia capillaris Artemisia scoparia Artemisia iwayomogi	□ 全草
金錢草	Lysimachiae Herba	앵초과	多年生 草本	통로황 廣金錢草	Lysimachia christinae Desmodium styracifolium	全草





어느 것이 독약일까요??





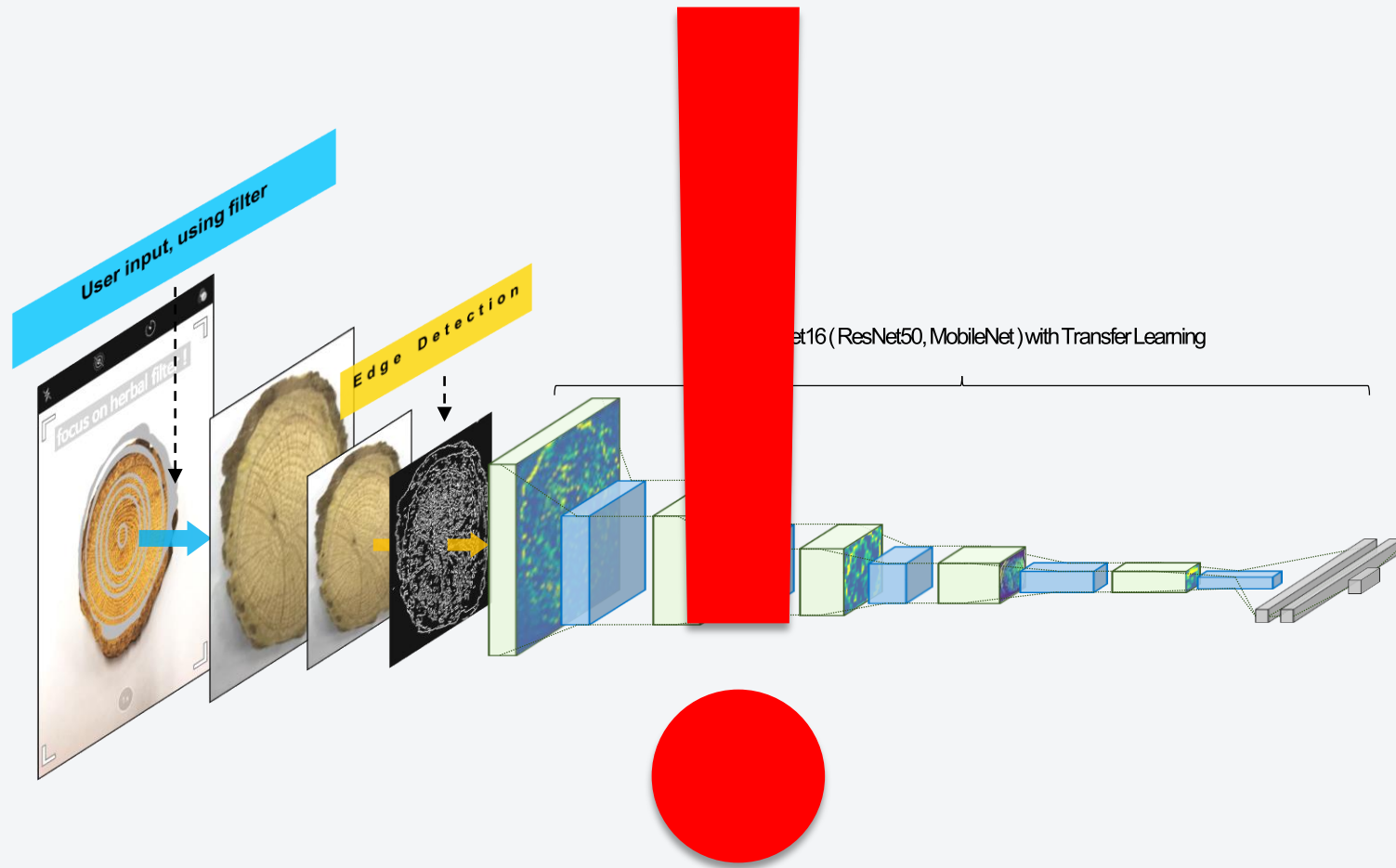
Toxic



네. 저도 잘 모릅니다.

독성 약재 감별은 전문가만이 가능한
영역이었으며, 현재의 약재 검증은
전문가가 manual로 직접 평가하는
시스템으로서 100개 중에 1개라도
놓친다면, 앞서 다뤘던 문제들을
발생하게 됩니다.







고려대학교
KOREA UNIVERSITY

빅데이터 기반의 지능 정보 시스템 개발 과정



추가 독성 약재 데이터 수집 불가



정보 시스템 개발 역량 부족



다양한 전공자들의 시너지효과로
정보 시스템으로 발전 방향 및 솔루션 논의

팀 단위, 약재 데이터 현황 파악 후
약재 데이터 추가 수집 요청

정보 시스템 개발 과정 학습 및 실습

데이터 요청 및 사전 조사, 수집, 촬영

데이터 수집에 앞서 문제점 파악



연구기관에 표준화된 약재 이미지 데이터가 **디지털화** 되어있지 않다.

표준화된 양식을 제시하면서 촬영을 해야한다.



독성 약재는 일반인이 **구매**할 수 없다.

연구기관이나 타 단체와의 협업이 필요하다.



독성 약재를 구하더라도, 그 약재와 감별해야 할 외형 **유사약재 리스트**가 필요하다.

전문가와 도메인 지식이 필요하다.

데이터 수집에 앞서 문제점 파악



연구기관에 표준화된 약재 이미지 데이터가 **디지털화** 되어있지 않다.

표준화된 양식을 제시하면서 촬영을 해야한다.



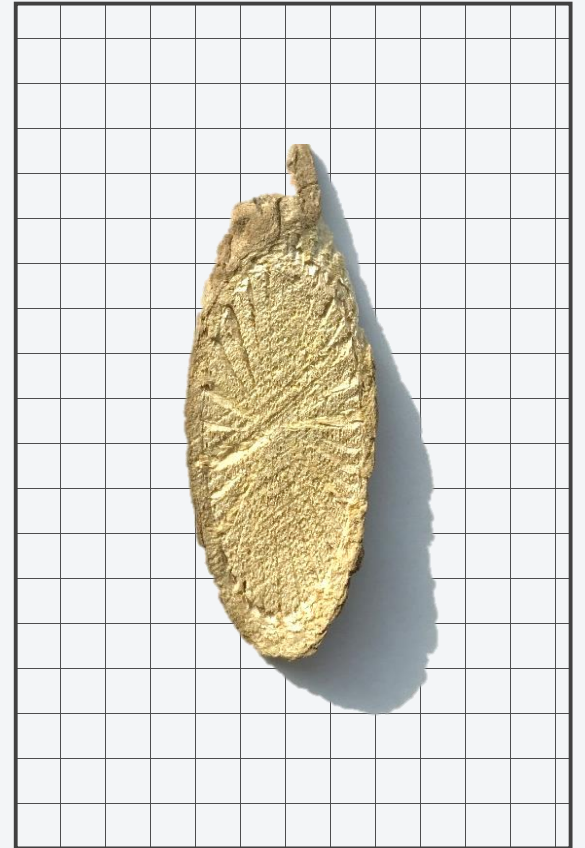
독성 약재는 일반인이 **구매**할 수 없다.

연구기관이나 타 단체와의 협업이 필요하다.

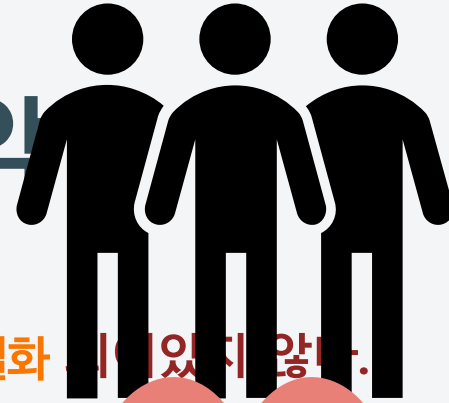


독성 약재를 구하더라도, 그 약재와 감별해야 할 외형 **유사 약재 리스트**가 필요하다.

전문가와 도메인 지식이 필요하다.



데이터 수집에 앞서 문제점 파악



연구기관에 표준화된 약재 이미지 데이터가 **디지털화** 되어 있지 않다.

표준화된 양식을 제시하면서 촬영을 해야 한다.



독성 약재는 일반인이 **구매**할 수 없다.

연구기관이나 타 단체와의 협업이 필요하다.



독성 약재를 구하더라도, 그 약재와 감별해야 할 외형이 **유사**하다. **구분** 필요하다.

전문가와 도메인 지식이 필요하다.



데이터 수집에 앞서 문제점 파악



연구기관에 표준화된 약재 이미지 데이터가 **디지털화** 되어있지 않다.

표준화된 양식을 제시하면서 촬영을 해야한다.

Toxic



독성 약재는 일반인이 **구매**할 수 없다.

연구기관이나 타 단체와의 협업이 필요하다.

Toxic



독성 약재를 구하더라도, 그 약재와 감별해야 할 외형 **유사 약재 리스트**가 필요하다.

전문가와 도메인 지식이 필요하다.

독약한방

약재 **이미지** 데이터 베이스



- 3개의 카테고리
- 19종의 약재
- 약재당 200장의 이미지 데이터

약재 **정보** 데이터 베이스



- 약재 기본 정보 데이터
- 독성약재 유사그룹 | 한방분류 그룹
- 최신 뉴스 | 최신 논문

1. 關木通, 木通, 青風藤, 粉防己
2. 茵陳蒿, 韓茵陳, 青蒿
3. 麻黃, 木賊
4. 牽牛子, 蘿菔子, 決明子, 枸杞子, 補骨脂
5. 芫花, 旋覆花, 金銀花, 甘菊
6. 白附子, 薤白
7. 甘遂, 大戟, 巴戟
8. 大黃, 肉蓯蓉
9. 杏仁, 桃仁



1. 단면류 : 關木通, 木通, 青風藤, 粉防己, 大黃, 肉蓯蓉
2. 줄기류 : 甘遂, 大戟, 巴戟
3. 씨앗류 : 牽牛子, 蘿菔子, 決明子, 枸杞子, 補骨脂, 杏仁, 桃仁



데이터 사전 조사 (인터넷, 전공서적 조사)을 통해
딥러닝 모델의 input으로 적절한 카테고리 분류

이승인교수님

장상재, 녹두, 박하, 육두구, ★연두, ★속수자, ★파두, ★마전자, 사인, 백두구, 차전자, 나복자, 빈랑자, 도인, 백개자, 동과자, ★행인, 소자, 정력자, ★백과, 산조인, 백자인, 결명자, 호로파, 보글지, 호도인, 토사자, 사원질려, 구자, 부소맥, 육두구, 연자육, 검실, ★대중자, 건을, ★급성자, 면실자, 육별자, ★보두, 지구자, 흑두, ★비마자

저장

종자류 - 우방자, 청상자, 녹두, 화마인, 육리인, ★견우자, ★속수자, ★파두, ★마전자, 사인, 백두구, 초두구, 의이인, 적소두, 차전자, 동규자, 여지핵, 나복자, 빈랑자, 비자, 무이, 왕불류형, 중을자, 백개자, 반대해, 동과자, ★행인, 소자, 정력자, ★백과, 산조인, 백자인, 결명자, 호로파, 보글지, 호도인, 토사자, 사원질려, 구자, 부소맥, 육두구, 연자육, 검실, ★대중자, 건을, ★급성자, 면실자, 육별자, ★보두, 지구자, 흑두, ★비마자

오전 11:00

본초학 전공 교수님의 도움으로 작성된
독성약재 및 외형이 비슷하여 감별이 필요한
약재 리스트 완료 작성

이승인교수님

실습실구비약재(23종) - 우방자, 육리인, 사인, 백두구, 차전자, 나복자, 빈랑자, 도인, 백개자, 동과자, ★행인, 소자, 정력자, 산조인, 백자인, 호로파, 보글지, 토사자, 구자, 육두구, 연자육, 검실,

미구비약재 - 청상자, 녹두, 화마인, ★견우자, ★속수자, ★파두, ★마전자, 초두구, 의이인, 적소두, 동규자, 여지핵, 비자, 무이, 왕불류형, 중을자, 반대해, ★백과, 결명자, 호로파, 보글지, 토사자, 구자, 부소맥, 육두구, 연자육, 검실, ★대중자, 건을, ★급성자, 면실자, 육별자, ★보두, 지구자, 흑두, ★비마자

이승인교수님

종자류 실습실구비약재 중 15종, 미구비약재 중 독성약재 5종으로 가자. 총 20종.

촬영도 깔끔하고, 인식률도 높을 것으로 예상되고

2019. 7. 22

연구기관 데이터 요청 및 현재 확보 데이터 파악 약초 전공 교수님과 협업

이승인교수님

산형과, 미나과 28종에서는 지금
2개(관엽강활, 대나물)을 못구해서
기다려야 함.

대나물은 오늘 여수에 가서 채집
가능한지 불 예정이고

관엽강활은 중국에 배송요청하 놓은
상태임

오전 10:33

대황은 단면있는것을 못구할까요?
뿌리단면으로 같것같은데요

오전 10:34

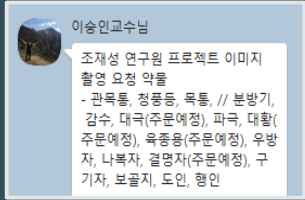
이승인교수님

대황은 단면은 맞는데 너무 커서
포개놓은 것임

오전 10:34

국내에서 구할 수 없는
독성, 일반 약재도
국제 배송 요청!!

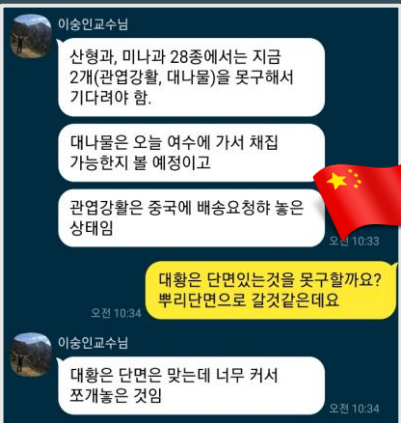




2019. 7. 22 다양한 논의를 통한 디지털화 약재 선정 및 연구기관 데이터 촬영 시작!

2019. 8. 1

연구기관 데이터 요청
및 현재 확보 데이터 파악
약초 전공 교수님과 협업

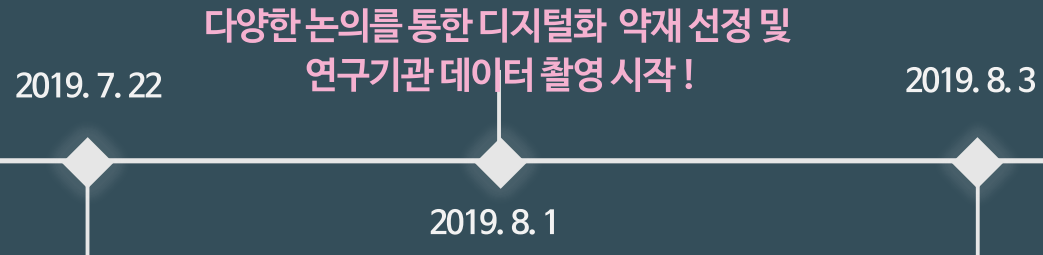
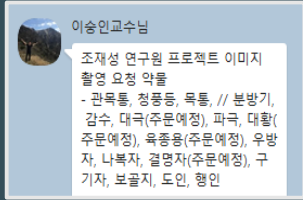


국내에서 구할 수 없는
독성, 일반 약재도
국제 배송 요청!!

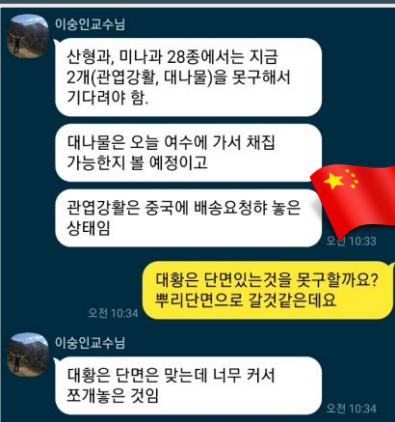


01 이미지데이터 와 정보데이터의 수집

데이터 수집



연구기관 데이터 요청
및 현재 확보 데이터 파악
약초 전공 교수님과 협업



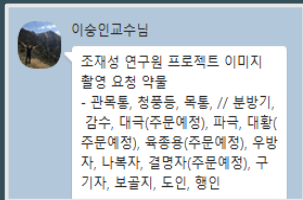
국내에서 구할 수 없는
독성, 일반 약재도
국제 배송 요청!!



프로젝트 팀원들의 직접촬영
연구실 인력 부족 문제 및
빠른 데이터 확보의 필요성



01 이미지데이터 와 정보데이터의 수집 데이터 수집



2019. 7. 22

다양한 논의를 통한 디지털화 약재 선정 및
연구기관 데이터 촬영 시작!

2019. 8. 3

딥러닝 1차 학습 시작!

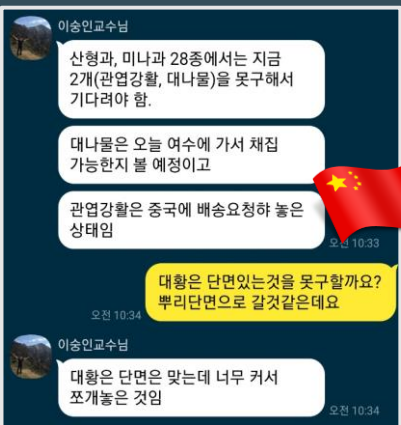
VggNet | ResNet | MobileNet

2019. 8. 5 ~ 8

연구기관 데이터 요청
및 현재 확보 데이터 파악
약초 전공 교수님과 협업

프로젝트 팀원들의 직접촬영

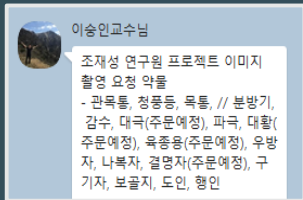
연구실 인력 부족 문제 및
빠른 데이터 확보의 필요성



국내에서 구할 수 없는
독성, 일반 약재도
국제 배송 요청!!



01 이미지데이터 와 정보데이터의 수집 데이터 수집

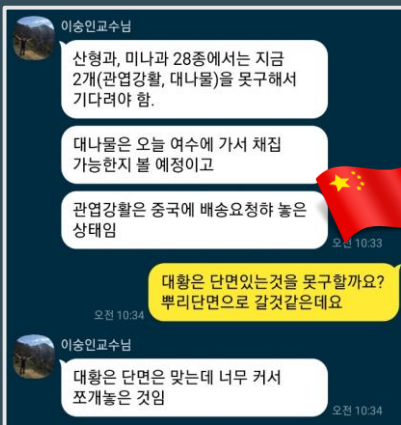


2019. 7. 22

다양한 논의를 통한 디지털화 약재 선정 및
연구기관 데이터 촬영 시작!

2019. 8. 1

연구기관 데이터 요청
및 현재 확보 데이터 파악
약초 전공 교수님과 협업



국내에서 구할 수 없는
독성, 일반 약재도
국제 배송 요청!!



2019. 8. 3

딥러닝 1차 학습 시작!

VggNet | ResNet | MobileNet

2019. 8. 5 ~ 8

프로젝트 팀원들의 직접촬영

연구실 인력 부족 문제 및
빠른 데이터 확보의 필요성



2019. 8. 9 ~ 10

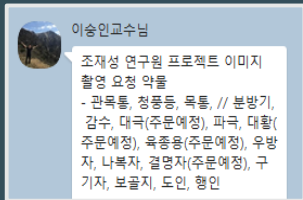
1차 모델 평가

정확도 | f1-score | ROC curve

1차 모델 히스토리 관리

엑셀 | 주피터 노트북





2차 적합 모델 조사

DenseNet
LeafNet
Attention Mecanism
MobileNet V2
AntNet

2019. 7. 22

다양한 논의를 통한 디지털화 약재 선정 및
연구기관 데이터 촬영 시작!

2019. 8. 3

딥러닝 1차 학습 시작!

VggNet | ResNet | MobileNet

2019. 8. 9 ~ 10

2019. 8. 1

2019. 8. 5 ~ 8

2019. 8. 11 ~

연구기관 데이터 요청
및 현재 확보 데이터 파악
약초 전공 교수님과 협업

프로젝트 팀원들의 직접촬영

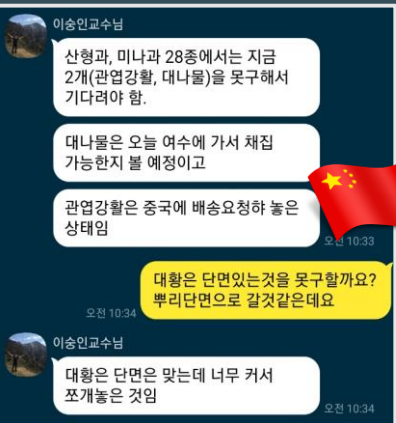
연구실 인력 부족 문제 및
빠른 데이터 확보의 필요성

1차 모델 평가

정확도 | f1-score | ROC curve

1차 모델 히스토리 관리

엑셀 | 주피터 노트북



국내에서 구할 수 없는
독성, 일반 약재도
국제 배송 요청!!



독약한방

약재 이미지 데이터 베이스



- 3개의 카테고리
- 19종의 약재
- 약재당 200장의 이미지 데이터

약재 정보 데이터 베이스



- 약재 기본 정보 데이터
- 독성약재 유사그룹 | 한방분류 그룹
- 최신 뉴스 | 최신 논문

01 이미지데이터 와 정보데이터의 수집 데이터 수집



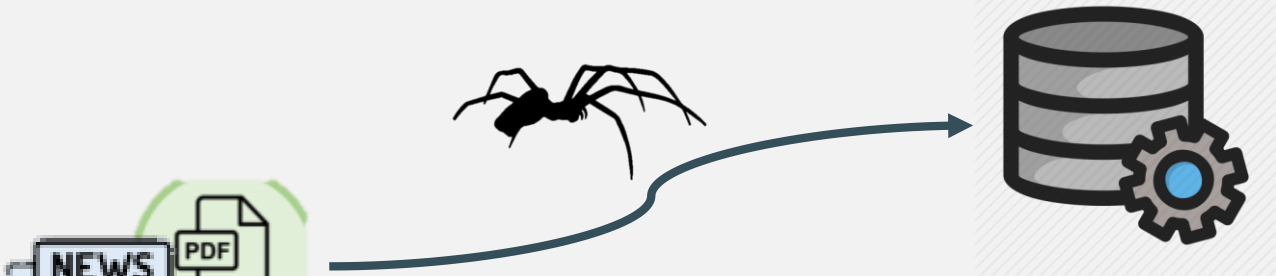
약재 기본 정보 데이터 최초 크롤링

이승인 교수님

- scientific name은 theplantlist에서 검토하고 수정해 줄게.
- figure-feature도 [본초학] 내용 정리해줄게.
- 유통되는 대활의 종류가 3종이고, 전부 독성이 있지. 따라서 3종 모두 감별해서 독성정보 제공하는 것이 의미가 있음. 독성정보는 3종이 유사함.
- 형태가 유사해서 헷갈려서 감별이 되는지 궁금하기도 함.
- 상록은 대활 용종용과 외관이 유사하고, 상록도 독성약재임. 빼고 되고, 안빼도 됨.
- 상록은 국내에서도 많이 자생하고 있음.
- 약재 위치 정보상으로 19종 중 절반 이상이 외래종일 것임. [본초학]에 산지 정보 나와 있음.

독성 정보 및 raw데이터의 전문가 검수

독약한방



해당 약재별 최신 뉴스, 최신 논문 크롤링

자체 리스트를 활용한 독성 및 유사약재의 그룹 관리

- 단면류: 關木通, 木通, 青風藤, 粉防己, 大黃, 肉蓯蓉
- 줄기류: 甘遂, 大戟, 巴戟
- 씨앗류: 牽牛子, 蘿藦子, 決明子, 枸杞子, 補骨脂, 杏仁, 桃仁

약재 정보 데이터의 ERD

Category					
한방분류 (Category)					
PK	AI	FK/Null	Logical Name	Name	Type
✓	✓	+	한방분류번호	category_id	INT
		+	한방분류명	category_name	VARCHAR(255)

Similarity_group					
유사약재그룹 (Similarity Group)					
PK	AI	FK/Null	Logical Name	Name	Type
✓	✓	+	그룹번호	group_id	INT
		+	그룹명	group_name	VARCHAR(255)

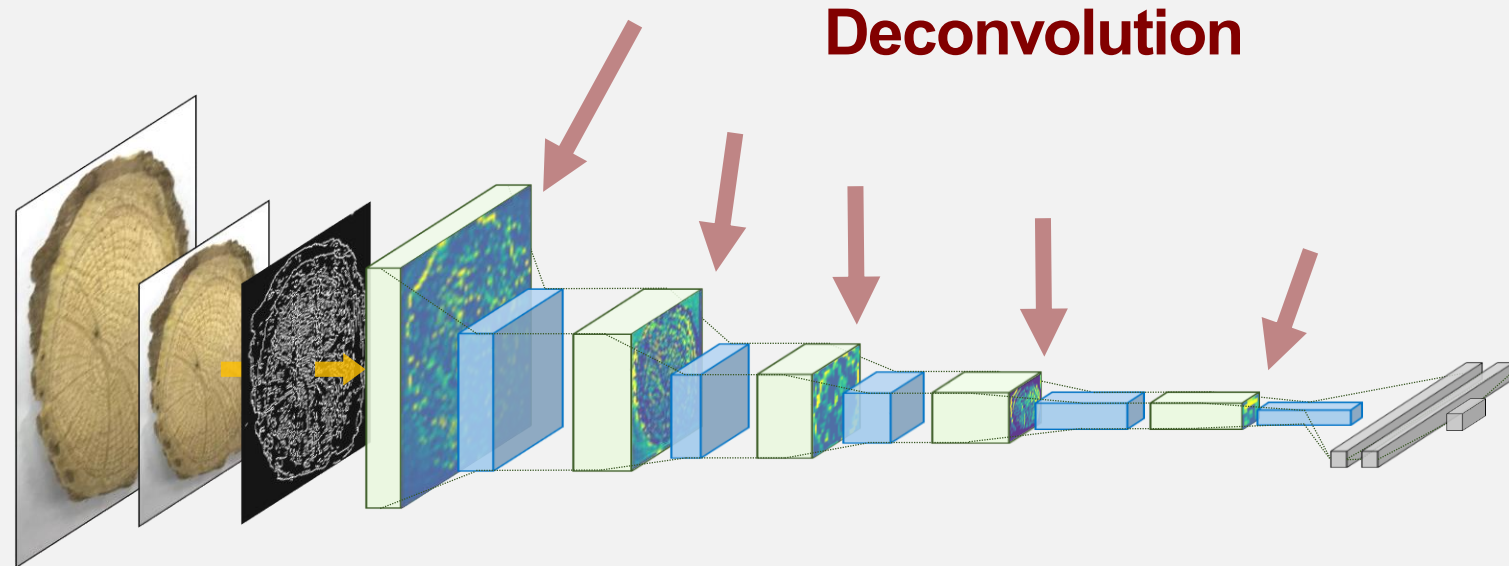
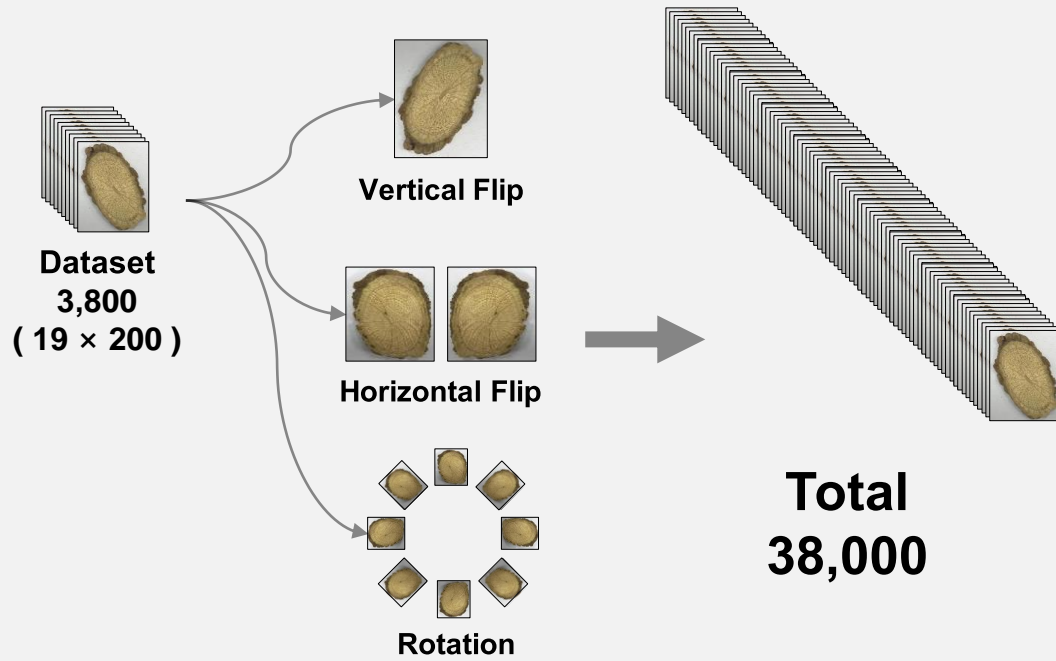
Location					
약재위치 (Location)					
Num	PK	AI	FK/Null	Logical Name	Name
1	✓	✓	+	위치번호	location_id
2			✓	약재번호	herb_id
3			+	국가명	country
4			+	위치(지명)명	location_name
5			+	위도	pos_x
6			+	경도	pos_y

Herb					
약재 (Herb)					
PK	AI	FK/Null	Logical Name	Name	Type
✓	✓	+	약재번호	herb_id	INT
	✓	✓	한방분류번호	category_id	INT
	✓	✓	유사약재그룹번호	group_id	INT
		+	약재명	herb_name	VARCHAR(100)
		+	주치	major_effect	VARCHAR(50)
		+	효능	effect	VARCHAR(255)
		+	기원	origin	VARCHAR(255)
		+	학명	scientific_name	VARCHAR(255)
		+	독성여부 (0 / 1)	is_poison	INT(1)
	✓	+	외관상특징	figure_feature	VARCHAR(3000)
	✓	+	이미지경로	image_path	VARCHAR(1000)

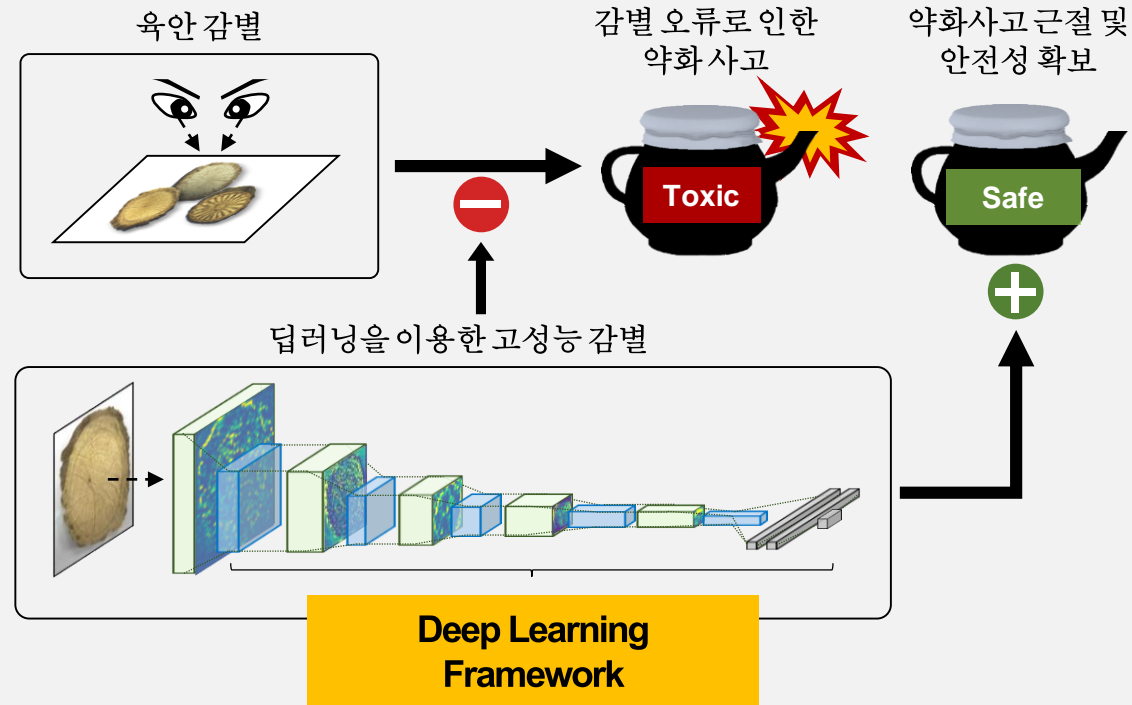
News					
뉴스					
PK	AI	FK/Null	Logical Name	Name	Type
✓	✓	+	뉴스번호	news_id	INT
	✓	+	약재번호	herb_id	INT
		+	뉴스제목	title	VARCHAR(3000)
		+	뉴스일자	date	DATETIME
		+	내용	content	VARCHAR(5000)
		+	URL	url	VARCHAR(255)

Journal					
논문					
PK	AI	FK/Null	Logical Name	Name	Type
✓	✓	+	논문번호	journal_id	INT
	✓	+	약재번호	herb_id	INT
		+	논문제목	title	VARCHAR(3000)
		+	논문일자	date	DATETIME
		+	URL	url	VARCHAR(255)

Augmentaion



Critical logic



✓

Model history

Basic Information				Model Information										Used Dataset Information			
Name	Accuracy (%)	Auucracry (count)	Result Analysis (Summary)	Base Model	Pre-trained	Batch_size	Epochs	Learning Rate	Optimizer	Loss(criterion)	Training_Loss	Validation_Loss	Model_Unique_Explanation (Summay)	Data	Train	Test	Validation
pure_vgg16	33		-	VGG16	FALSE	4	1	0.001	Adam	Cross_entropy		-	-	폴기	480	120	0
pre_vgg16	95		- Pre-trained의 성능이 훌륭하다.	VGG16	TRUE	4	1	0.001	Adam	Cross_entropy		-	-	폴기	480	120	0
fixed_vgg16	34		- General purpose pre-trained weights should be updated for specific dataset.	VGG16	TRUE	4	1	0.001	Adam	Cross_entropy		-	- feature extraction 부의 pre-trained 파라미터를 Fixed로 설정	폴기	480	120	0
resnet_pre-traine	90		-	ResNet50	TRUE	4	5	0.001	Adam	Cross_entropy		-	-	씨앗A	800	198	0
pre_resnet18	95	119/120	- Pre-trained의 성능이 훌륭하다.	ResNet18	TRUE	4	5	0.001	Adam	Cross_entropy	0.053	-		폴기	480	120	0
fixed_resnet18	98	118/120	- fixed임에도 불구하고 resnet은 good	ResNet18	FALSE	4	5	0.001	Adam	Cross_entropy	0.096	-	- feature extraction 부의 pre-trained 파라미터를 Fixed로 설정	폴기	480	120	0
fixed_vgg16	92	184/200	- fixed임에도 불구하고 단면은 good	VGG16	FALSE	4	5	0.001	Adam	Cross_entropy		-		단면A	800	200	0
pre_resnet18	98	626/638		ResNet18	TRUE	4	5	0.001	Adam	Cross_entropy	0.009	-		Total	2560	638	0

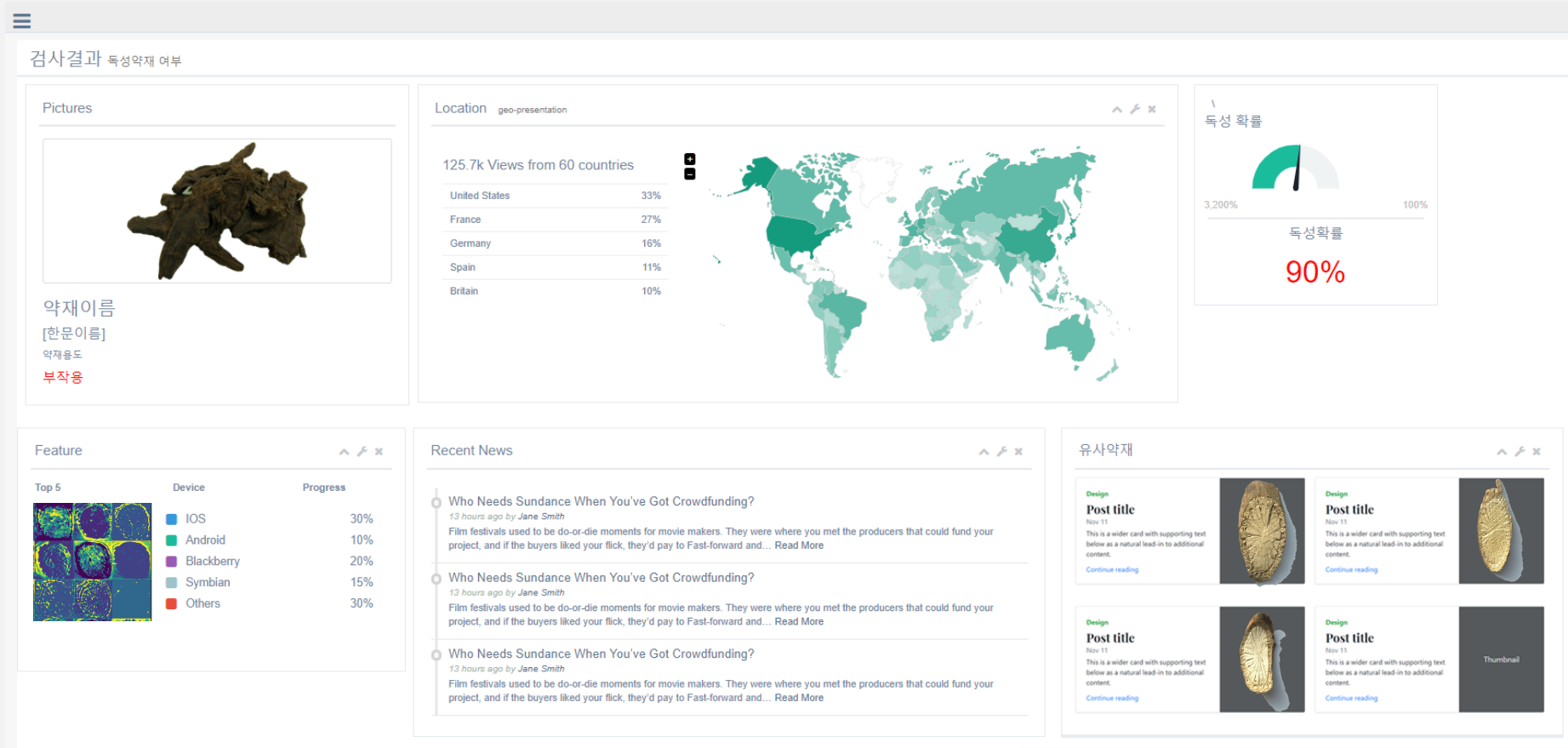
실현 가능한
독성 약재 감별 및
약재 정보 제공 시스템

독약한방



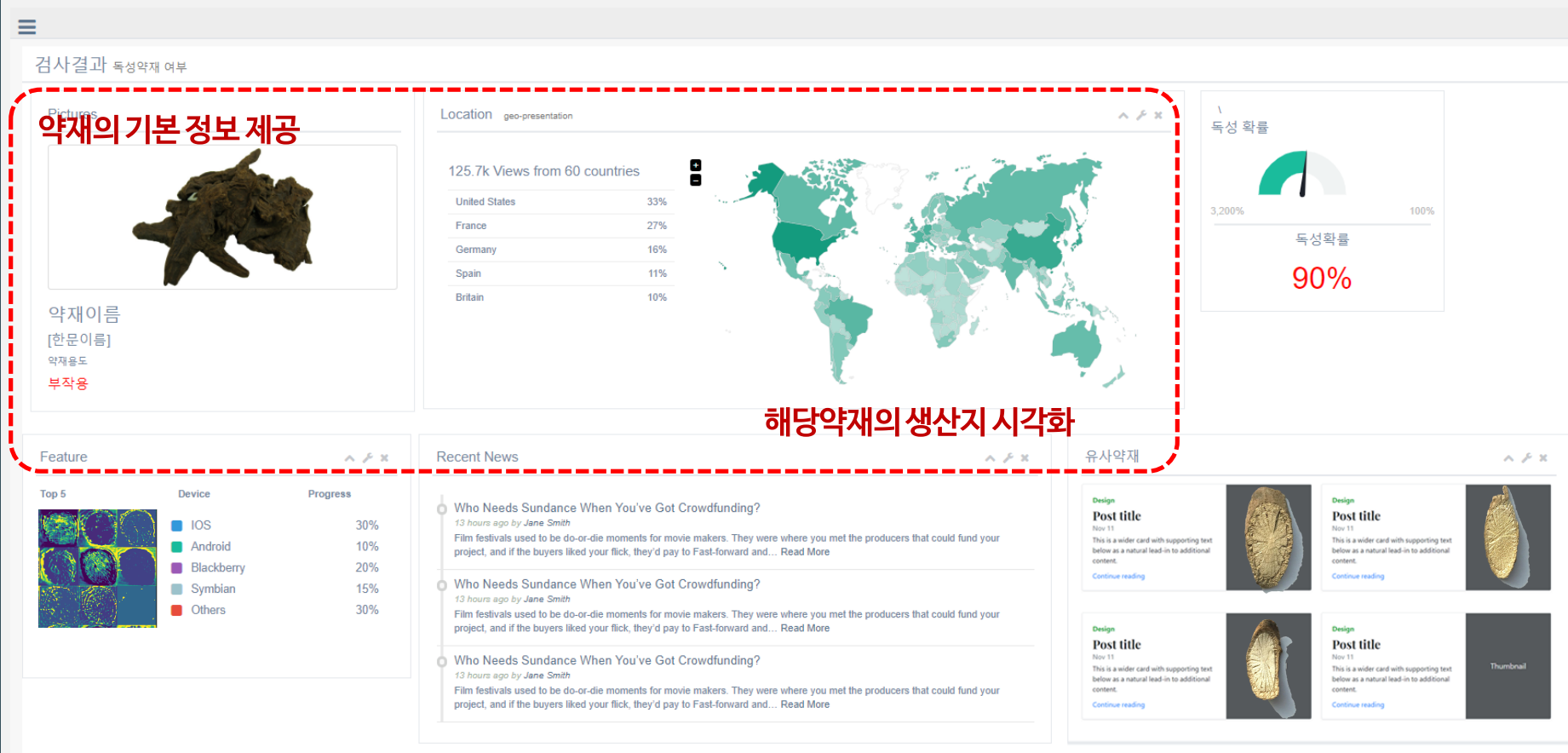
실현 가능한 독성 약재 감별 및 약재 정보 제공 시스템

독약한방



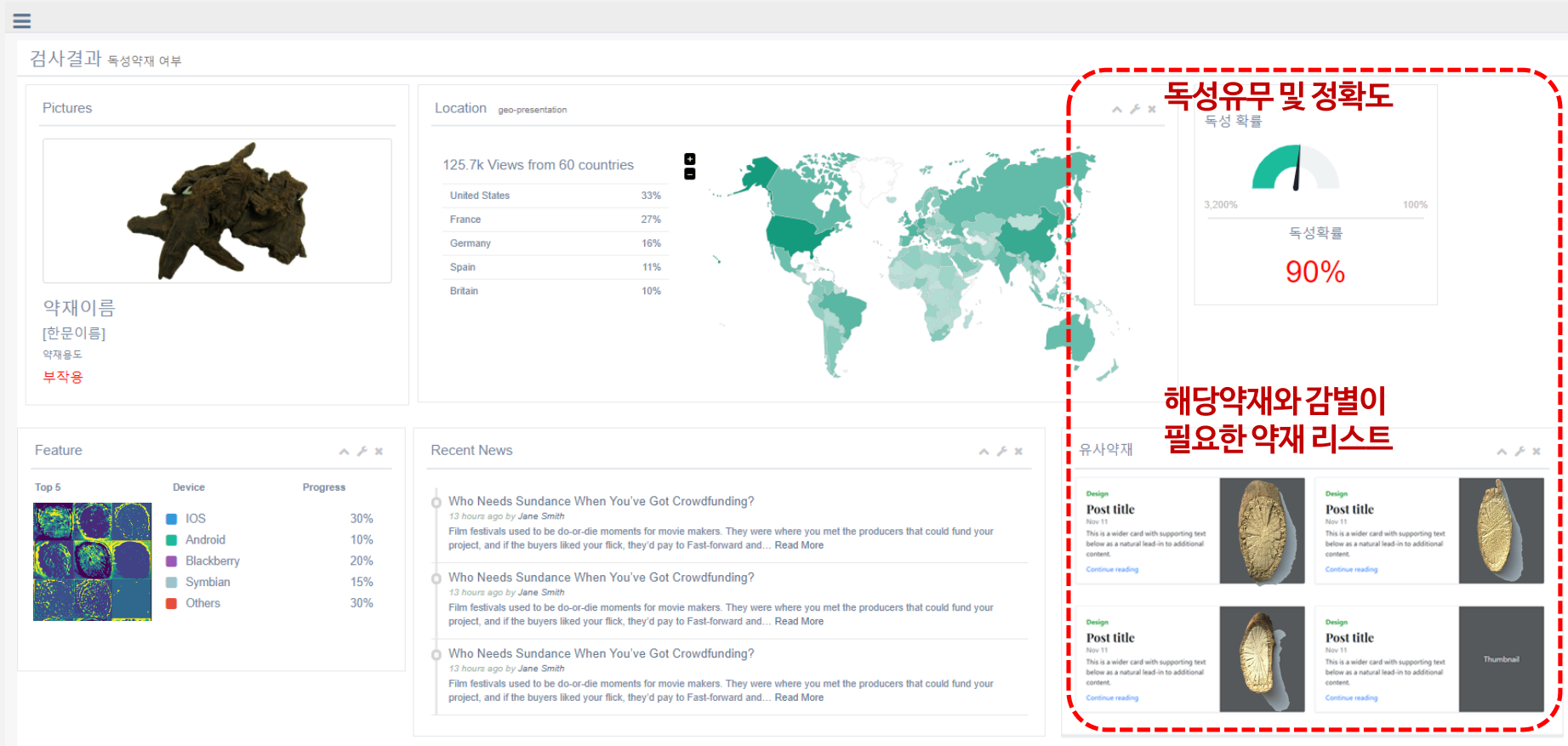
실현 가능한 독성 약재 감별 및 약재 정보 제공 시스템

독약한방



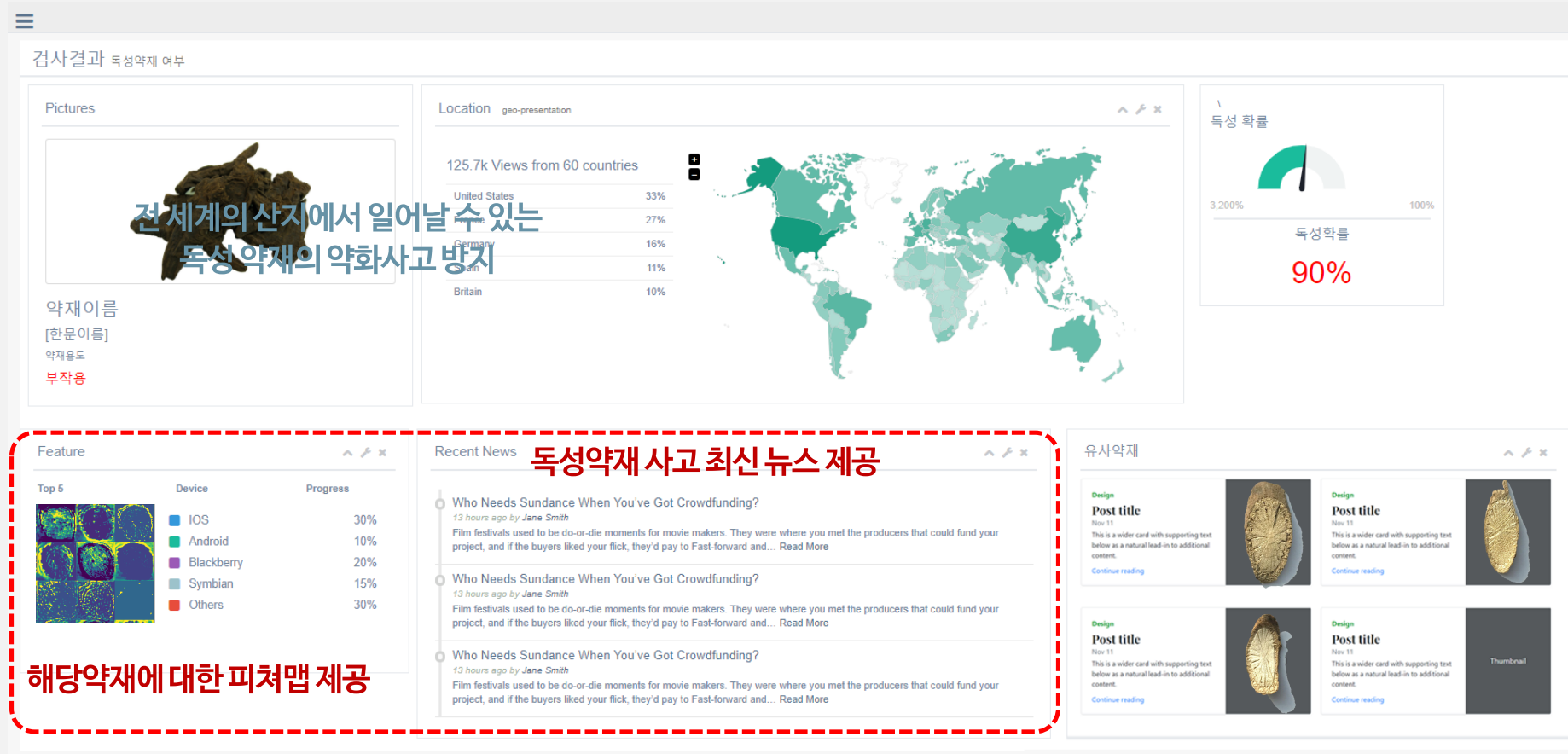
실현 가능한 독성 약재 감별 및 약재 정보 제공 시스템

독약한방



실현 가능한 독성 약재 감별 및 약재 정보 제공 시스템

독약한방



실현 가능한
독성 약재 감별 및
약재 정보 제공 시스템

독약 **한방**



Global 일반인용

전 세계의 산지에서 일어날 수 있는
독성 약재의 약화사고 방지



전문가 양성 기관

약대, 한의학과, 한약학과, 산약회 등의
교육자료로 활용된다.



확장 가능성

개인 3종 -> 19종에 제한된 상황
-> 프로젝트의 중요성이 대두된다면, 모든 약재 정보를 수록



식약처 등 공공기관, 연구기관
전문가 보조 어플리케이션

전문가의 1%를 채워
독약으로 인한 사고를 방지한다.

APPENDIX

모델 관리 google drive

<https://drive.google.com/open?id=17gSullA-w7Ztcq2NdoJDkw06IGK95AmbvQ1e98I-8v4>

Attention Model

Residual Attention Network http://openaccess.thecvf.com/content_cvpr_2017/papers/Wang_Residual_Attention_Network_CVPR_2017_paper.pdf

CBAM

<https://arxiv.org/abs/1807.06521>

한의학에서의 AI 관심도 : 기사한의사들 "우리가 봐도 한방 표준화 멀었다."

<https://www.doctorsnews.co.kr/news/articleView.html?idxno=114660>

데이터 수집 과정

- 허브허브 (<https://herbhub.com/>)
- 식품의약품안전처 한약(생약) DB (http://www.nifds.go.kr/herb/m_442/list.do)
- 한의약융합연구정보센터 생약 up-to-date (https://www.kmcric.com/database/herb_search_list)

참고 문헌

- 본초감별도감 1-3권 (<https://drive.google.com/open?id=11zfoR5OHAPzDiYU9ulp-4orKUxfejE9z>)
- 알기쉬운 한약재감별법 (식품의약품안전청 발간,
https://drive.google.com/open?id=1bqehJn80ztn2Q_b9DVkiviR6CpVBVDQR))

데이터 사진 구글 드라이브

https://drive.google.com/drive/folders/1KKi-CrHmw2iV7V-_h8PcqLqQLZRvX7Fk

모델 weight

<https://drive.google.com/open?id=1Jvaqd38lhYeRxW9aPSnFvf35ITGwS4A8>

모델 모음

<https://drive.google.com/drive/folders/1Jvaqd38lhYeRxW9aPSnFvf35ITGwS4A8?zx=nxcwvlofpcxp>

독약한방

감사합니다