

# 동의보감 독초 판별 AI 학습용 이미지 데이터 가공 매뉴얼 Ver 2.0

(2021.09)



유한회사 나인펄스

☐ 제•개정 이력

버전	제•개정 일자	내용	작성	검토/승인
Ver 1.0	2021. 07. 25	최초버전	송운영	노인규
Ver 1.1	2021.08.13	내용 수정	송운영	노인규
Ver 1.2	2021.08.19	Labelme 사용방법 추가	송운영	노인규
Ver 2.0	2021.09.06	허브넷 가입 & 사용방법 추가	송운영	노인규

## 본 매뉴얼은

한국지능정보사회진흥원(NIA) 데이터 댐 구축(2차)사업 추가과제로 선정된

‘동의보감 독초 이미지 데이터 구축’ 수행을 위한 **데이터 가공 매뉴얼**입니다

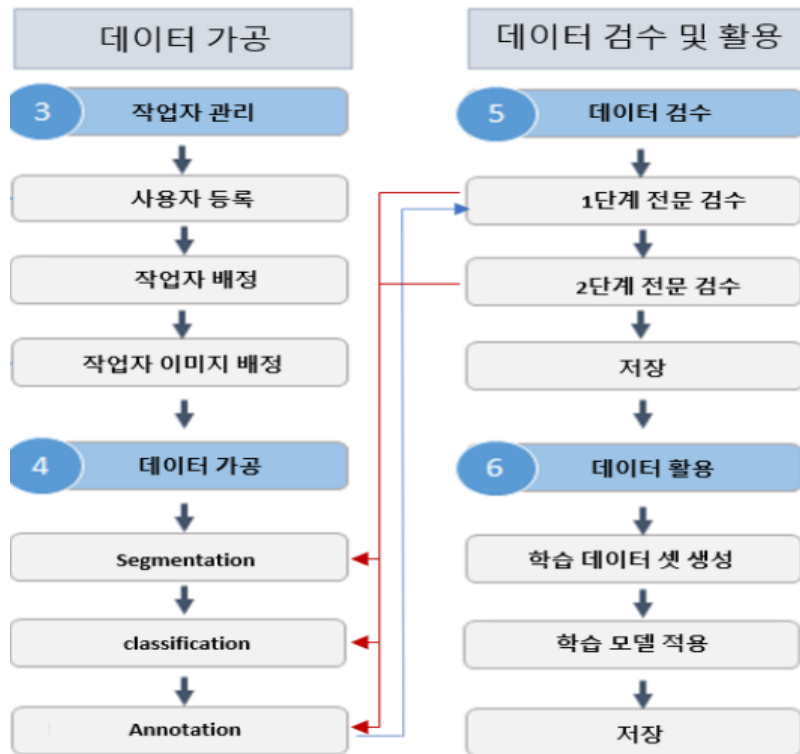
- 01 데이터 가공
- 02 데이터 가공도구 [허브넷]
- 03 데이터 가공 작업(Labelme for Win)
- 04 결과물

# CONTENTS

## 가공 목적

**가공 목적**, 작업자에게 필요한 지침을 제공하고 작업 간에 발생하는 시간과 비용을 축소하기 위한 목적

### 가공 기준 및 처리 절차



.가공(Labeling)->어노테이션(Annotation)

->저장 프로세스에서 학습 조건에 맞는 가공(Labeling) 방법 및

어노테이션(Annotation)을 설정하고, 인공지능 학습데이터에 대한 공통 기준으로 활용

.데이터 가공(Labeling)은 학습에 필요한 어노테이션(Annotation) 정보기준을 설정하고 가공(Labeling)시 적용

.학습데이터는 예측하고자 하는 데이터와 가장 유사해야 함

.클라우드워커에 개인 업무 배정 시 작업자별로 폴더 권한을 부여하여 가공(Labeling) 작업의 효율성 및 데이터의 중복 작업 방지

.가공(Labeling) 매뉴얼에 따라 가공(Labeling) 작업 수행 및 어노테이션(Annotation)

.가공(Labeling) 작업 완료 후 클라우드워커 간 교차 검수

.디자인 전문가의 2차 가공(Labeling) 검수



## 가공 단계

### 가공단계 ○ 라벨링·어노테이션(Annotation)작업방법

- 식별 가능한 객체를 대상으로 라벨링 작업 시행
- 라벨링 작업 시 객체의 분류(class)가 올바른지 확인  
(판단이 어려운 경우 필수로 담당자에게 확인 후 진행)
- 객체의 분류(class)마다 영역에 맞게 작업이 이루어졌는지 확인
- 약초, 유사 식물 이미지에서 전초, 잎, 꽃, 열매를 바운딩박스, 폴리곤 형태로 영역을 지정한 후 어떤 부위인지 어노테이션(Annotation)
- 초점이 맞아 선명한 부분만 라벨링 실시
- 전초의 경우 전체 부위를 영역으로 저장
- 라벨링 객체 외곽 기준은 2픽셀 이내로 최대한 붙여서 영역을 지정
- 전초, 잎, 꽃, 열매 객체의 라벨링 영역이 70% 이상 되도록 지정
- 이미지가 겹치거나 가려진 객체는 라벨링을 하지 않음
- 이미지 당 최대 5개의 객체를 라벨링 함

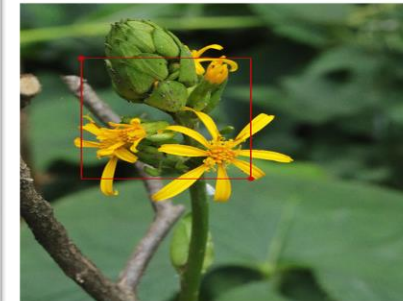
#### 가공 기준 및 방법

1. 사진의 식물에서 전체, 잎, 꽃, 열매를 BOX, 폴리곤 형태로 영역을 지정한 후 어떤 부분인지 지정
2. 포커스가 맞아 선명한 부분만 어노테이션 함
3. 대상 식물의 전체 부위를 영역으로 지정
4. 대상 식물의 외곽 기준은 2픽셀 이내로 최대한 붙여서 영역을 지정
5. 동일한 부분과 겹치거나 가려진 객체는 영역을 지정하지 않음
6. 사진 1장 당 최대 5개의 객체를 지정할 수 있음

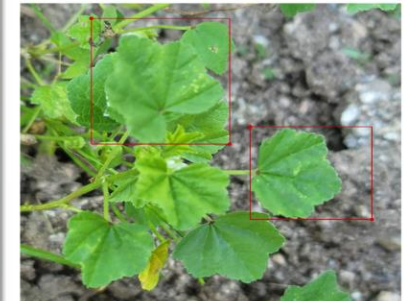
#### 올바른 예시



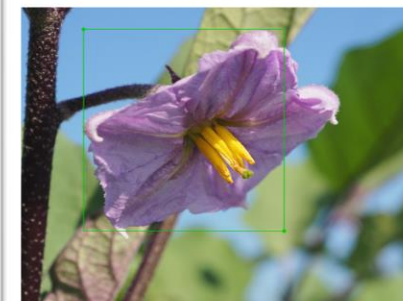
#### 잘못된 예시



전체 객체를 지정하지 않음



객체를 너무 넓게 지정함



객체를 벗어난 영역을 지정함



다른 객체를 지정함

# 1. 데이터 가공

## 가공 단계

## 가공 단계



(a) 원천 데이터



식물명	꽃물
약초명	하고초
어노테이션	flower
라벨링 위치	좌표 값

META DATA
카메라 속성 촬영 날짜, 시간, 제품명, 화소, 프레시, 워드/경도.....
촬영 지역(권역)
저작권 등의 정보

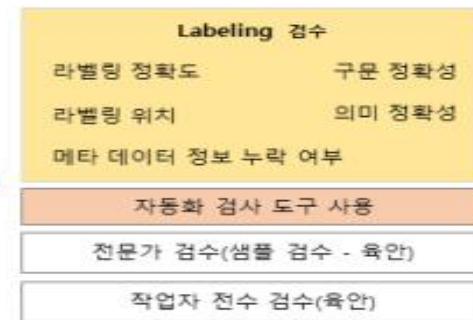
(c) Meta data



(g) 저장(폴더 구조 - 식물 이름 > 부위 > 촬영시기 > 라벨링 정보)



(f) 검수 결과 조치



(e) 육안 검수(작업자 교차 검수)



(c) Json



## 2. 데이터 가공도구 [허브넷]

### 가공 도구 [허브넷]



회원가입

로그인

과학기술정보통신부, NIA(한국정보화진흥원)가 추진하는 '인공지능 학습용 데이터 구축 사업(2차)' 선정과제

## 동의보감 약초판별 인공지능 학습용 데이터 구축

'동의보감 약초 판별 인공지능 학습용 데이터 구축'은

매년 독초를 약초로 오인, 남용하여 발생하는 독초에 대한 중독 및 사망 사고를

예방할 수 있는 약초와 독초 판별 인공지능서비스와

전문 지식 없이 약초를 과다 복용하거나, 무조건 좋다는 생각으로 음용하여

발생되고 있는 피해를 줄이기 위한 정확한 약초 사용 방법을 제공하기 위해

동의보감 수록 457종 중 올해 선정한 약초 50종, 비교식물 82종의 이미지 데이터와

정확한 약초 사용 방법을 위한 약초 정보를 구축하는 사업입니다.



### 데이터 가공은

허브넷([www.herbnet.kr](http://www.herbnet.kr))을 통해 진행되며

가공 작업에 참여하는 참여자는 허브넷에 가입하여야 한다.

가입 한 후 가공 책임자에게 아이디(허브넷 가입 이메일)를 알려주고 작업을 배정받는다

### 가공 작업 배정은

품목(약초) → 가공 단계(1차, 2차, 3차) 순으로 배정된다

가공 방법은 매뉴얼에 따라 수행하면 되고 궁금한 사항은 단체 채팅방을 통해 선임자들에게 요령을 익힌다

## 2. 데이터 가공도구 [허브넷]

### 플랫폼 개요

#### 1. 목적

약초(독초) 유사식물 이미지 데이터 수집, 데이터 정제, 데이터 가공, 학습용 데이터 변환 등의 업무를 위해 개발된 클라우드 소싱 참여 및 참여 대가 지급 등의 업무를 원활하게 지원하기 위하여 구축함

#### 2. 주소 : [www.herbnet.kr](http://www.herbnet.kr)(허브넷)

#### 3. 주요 기능

- ☐ 약초 및 비교식물 이미지 데이터 업로드
- ☐ 업로드 이미지 데이터에 대한 데이터 정제
  - 1차 : 품질 정제 / 2차 : 분류 정제 / 3차 : 범위 정제 / 4차 : 식물 분류 및 한의학 전문가 검수
- ☐ 클라우드 소싱
  - 클라우드 소싱 참여
  - 포인트 지급 및 확인
- ☐ 사용자 관리, 통계 관리, 수집 관리, 데이터 관리 등 플랫폼 관리

#### 4. 저작권 동의 : 저작물(촬영 이미지)에 대한 법적 사용 동의 절차 수행



## 2. 데이터 가공도구 [허브넷]

### [허브넷 회원가입]

#### 회원가입

아이디(이메일) \* ①

중복확인

비밀번호 \* ②

비밀번호 확인 \*

비밀번호는 영문 소문자, 숫자를 포함하여 8자리 이상으로 구성되어야 합니다.

이름 \* ③

생년월일 \* ④

성별 \* ⑤

☒ 남 ☐ 여

회원 사진 ⑥

파일 선택

⑦ ☐ 이용약관 및 개인정보처리방침에 동의합니다.

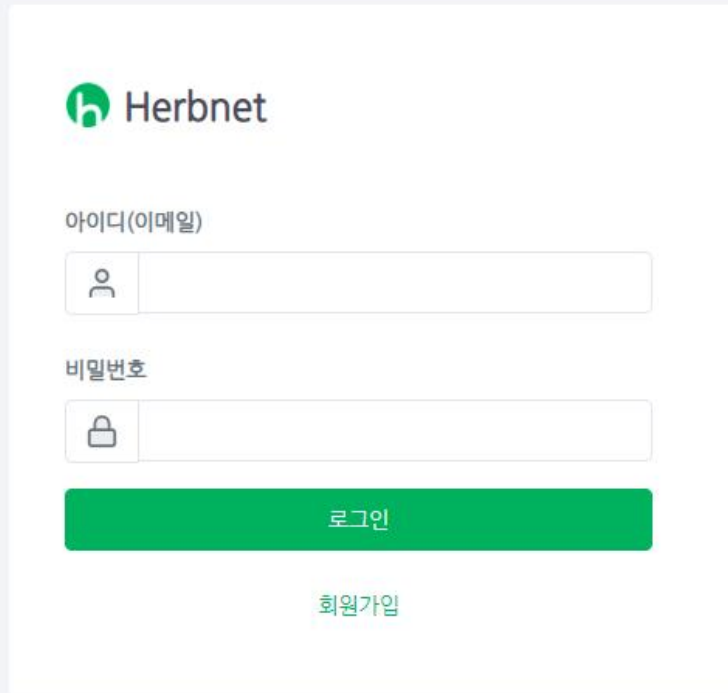
⑧ 가입하기

이미 계정이 있으신가요? 로그인하기 ⑨

- ① **아이디(이메일)**  
가입 이메일을 중복검사후 아이디로 설정합니다.
- ② **비밀번호**  
영문 소문자, 숫자를 혼합하여 8자리 이상으로  
구성되어야 합니다. 다시 비밀번호를 확인 합니다.
- ③ **이름**  
실명으로 입력합니다.
- ④ **생년월일**  
2000년 1월 1일인 경우, 20000101로 입력하시면 됩니다.
- ⑤ **성별** 성별을 체크합니다.
- ⑥ **회원사진**  
필수 사항이 아닌 선택 사항입니다.
- ⑦ **이용약관 및 개인정보처리방침에 동의합니다.**  
허브넷 가입을 위한 필수 사항입니다. 체크하시면 됩  
니다.
- ⑧ **가입하기**  
필수 입력사항을 모두 기입하고 가입하기 누르시면  
허브넷 회원이 됩니다.
- ⑨ **로그인하기**  
이미 계정이 있으시면 여기서 로그인 합니다.

## 2. 데이터 가공도구 [허브넷]

### 허브넷 로그인



The image shows the Herbnet login page. On the left, there is a white login form with the Herbnet logo at the top. Below the logo, there are two input fields: one for '아이디(이메일)' (ID/Email) and one for '비밀번호' (Password). Each field has a small icon (a person for ID, a lock for password) to its left. Below the password field is a green '로그인' (Login) button. At the bottom of the form is a link for '회원가입' (Sign Up). To the right of the form is a large image of green leaves. Overlaid on the bottom of this image is text in Korean: '동의보감 약초 이미지 학습 시데이터' (Ginseng Herb Image Learning Dataset), '약초 데이터가 쌓이면' (When herb data is accumulated), and '더 똑똑한 동의보감 약초 판별 인공지능이 됩니다.' (The Ginseng Herb discrimination AI becomes smarter).

로그인

허브넷([www.herbnet.kr](http://www.herbnet.kr))에 접속 후  
아이디(가입 이메일)와 비밀번호 입력  
접속

## 2. 데이터 가공도구 [허브넷]

### 가공 작업 1

Herbnet

2021년 09월 06일(월)

김환성  
검수자

데이터

사진등록

내가 등록한 사진

가공

라벨이미지 다운로드

json 파일 등록

커뮤니티

공지사항

문고 답하기

라벨작업용 이미지다운로드

약초 이름 (식물 이름)

검색

사진 용량문제로 다운로드 클릭시 20장씩 다운로드됩니다

식물명	수집부위	최종정제완료 건수	라벨완료 건수	남은 라벨건수	다운로드 완료 건수	다운로드 안한 건수	다운로드
가죽나무	전초 및 지상부	282	0	282	80	202	이미지 다운로드
		32	0	32	16	16	이미지 다운로드
		1,034	0	1,034	20	1,014	이미지 다운로드
		1,346	0	1,346	20	1,326	이미지 다운로드

### 가공 작업 1

① 라벨이미지 다운로드메뉴를 클릭

② 이미지 다운로드를 클릭하여 본인이 받고자 하는 파일을 다운로드

## 2. 데이터 가공도구 [허브넷]

### 가공 작업 2

동의보감 식물

라벨교육자료

라벨링자료

OneDrive

내 PC

3D 개체

다운로드

동영상

문서

바탕 화면

사진

음악

Windows (C:)

NP\_Thunbolt (F:)

RILLY (I:)

NP\_Thunbolt (F:)

이름

수정한 날짜

유형

가지-잎(11)

감국-잎17

공취-꽃(10)

공취-잎(11)

꽃물

꽃물-잎(10일)

난쟁이 아욱-잎(15)

능소화-꽃

능소화-잎

단삼-꽃

단삼-잎

도개비 가지-잎2

돼지감자-잎

마타리-잎

메밀-잎

무화과-열매

2021-09-11 오후 3:20

2021-09-09 오후 7:34

2021-09-09 오후 7:27

2021-09-10 오후 4:25

2021-09-17 오후 3:07

2021-09-10 오후 6:37

2021-09-20 오전 11:43

2021-09-12 오후 6:07

2021-09-10 오전 10:54

2021-09-15 오후 2:52

2021-09-20 오전 11:53

2021-09-11 오후 11:26

2021-09-22 오전 10:42

2021-09-10 오후 7:53

2021-09-15 오후 10:35

2021-09-10 오후 7:48

2021-09-10 오후 7:55

2021-09-15 오후 5:28

2021-09-12 오후 11:52

파일 폴더

파일 폴더

파일 폴더

파일 폴더

파일 폴더

파일 폴더

파일 폴더

파일 폴더

파일 폴더

파일 폴더

파일 폴더

파일 폴더

파일 폴더

파일 폴더

파일 폴더

파일 폴더

파일 폴더

파일 폴더

파일 폴더

### 가공 작업 2

③ 다운로드 받은 약초를 파일을 작업하기  
편리하게 분류



### 3. 데이터 가공 작업 [Labelme]

#### 가공 작업 3

④



Labelme.exe

Double click this icon

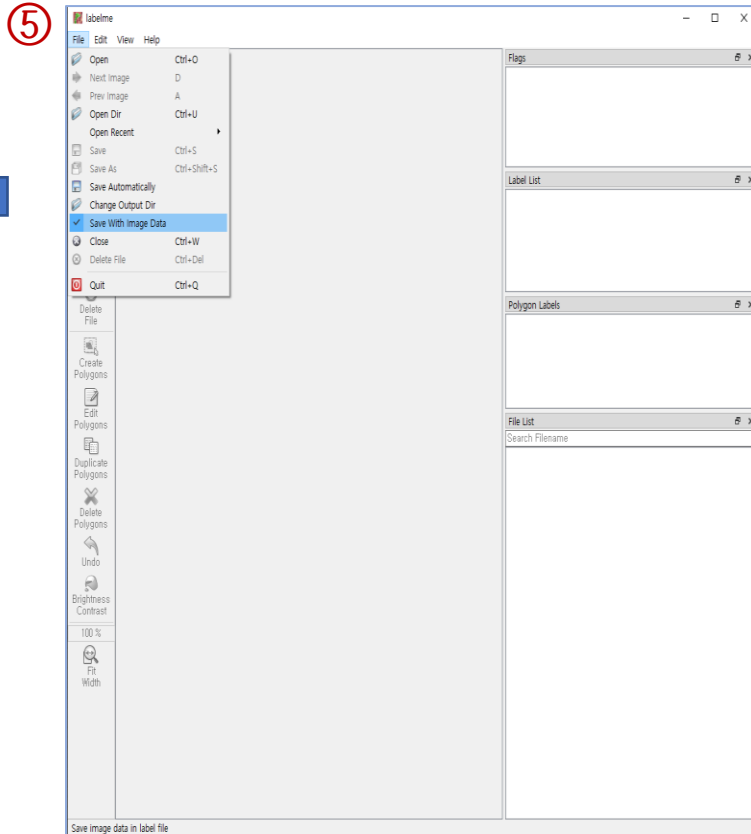
#### 가공 작업 3

④ 가공작업을 시작하기 위해

Labelme.exe를 클릭

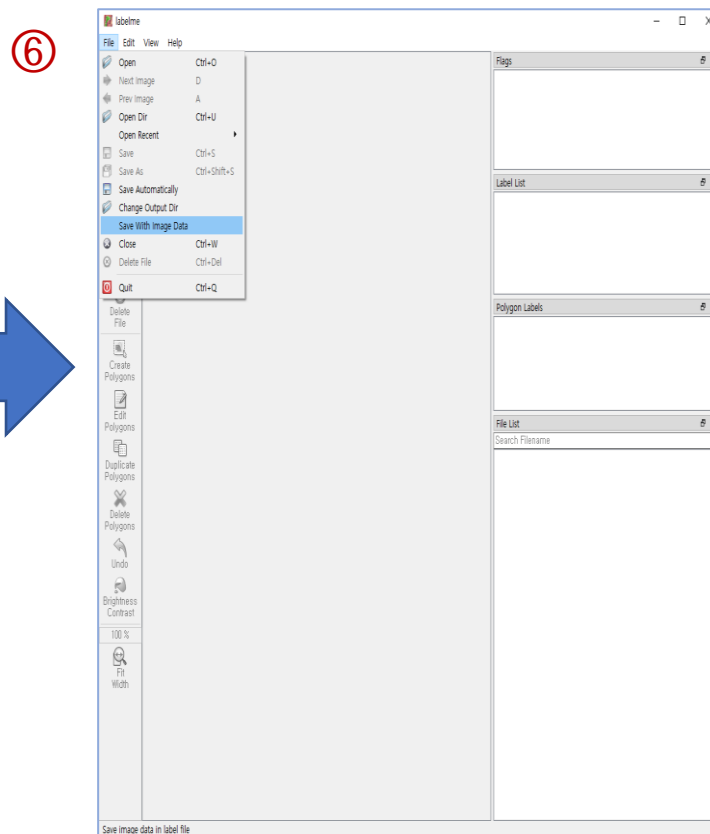
### 3. 데이터 가공 작업 [Labelme]

#### 가공 작업 4



Save with Image Data

=>



Save with Image Data

#### 가공 작업 4

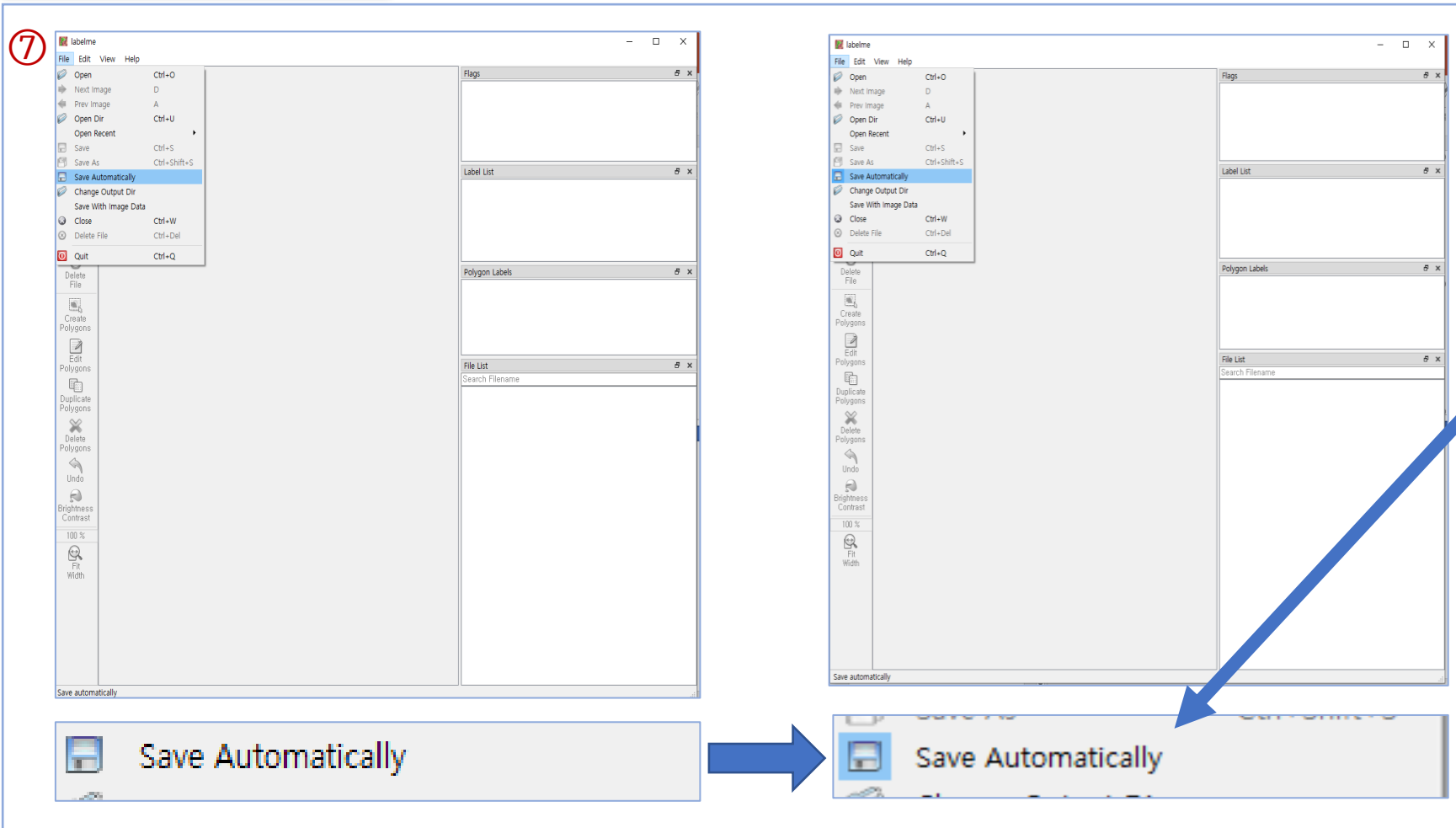
⑤ 작업전에 반드시 File에서  
Save with Image Data



⑥ ✓ 클릭하여 해제한다.

### 3. 데이터 가공 작업 [Labelme]

#### 가공 작업 5



#### 가공 작업 5

File에서 Save with Image Data를  
체크 후

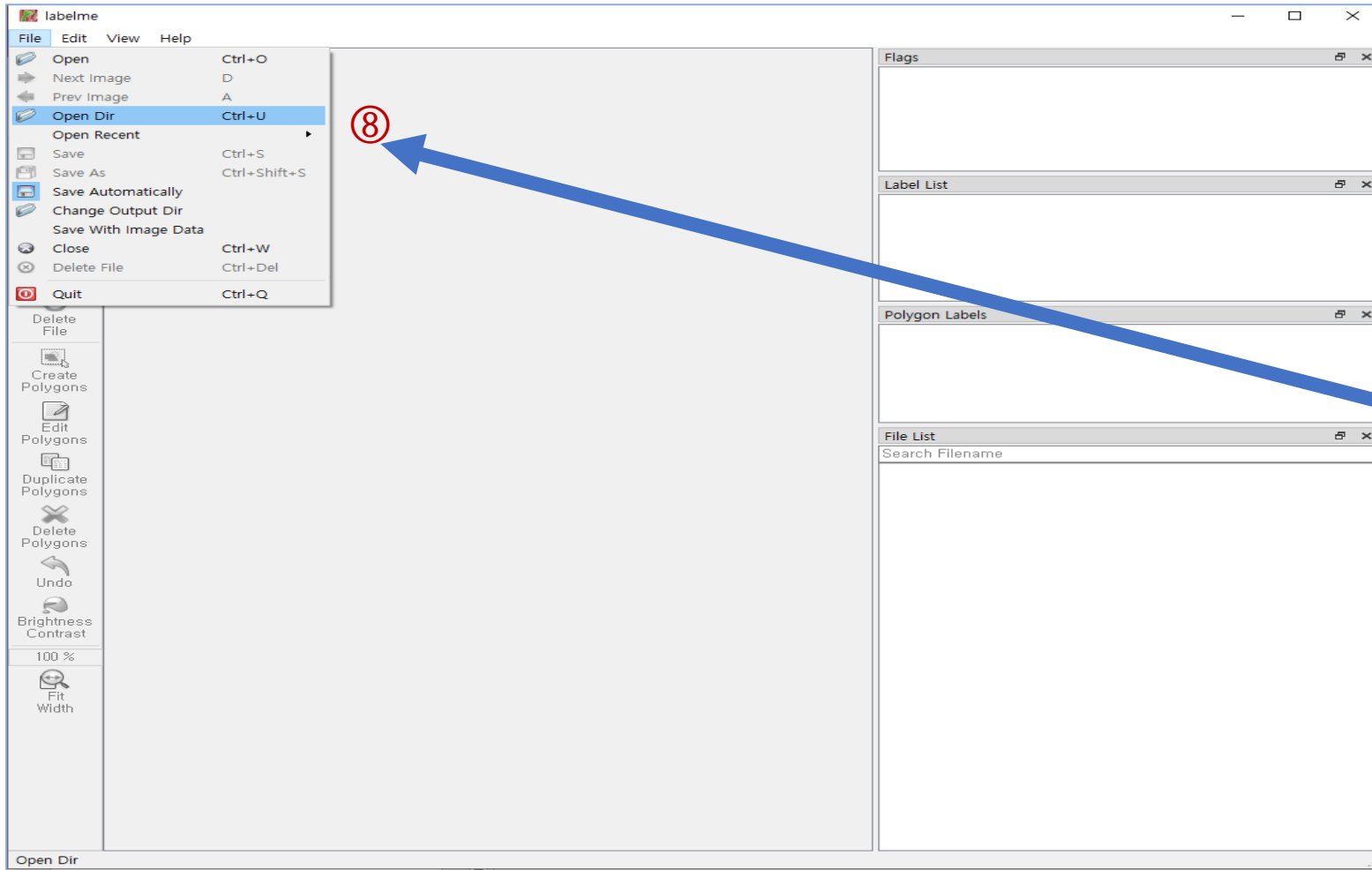
⑦ save Automatically클릭 후

작업을 진행해야

**자동으로 저장**

### 3. 데이터 가공 작업 [Labelme]

#### 가공 작업 6



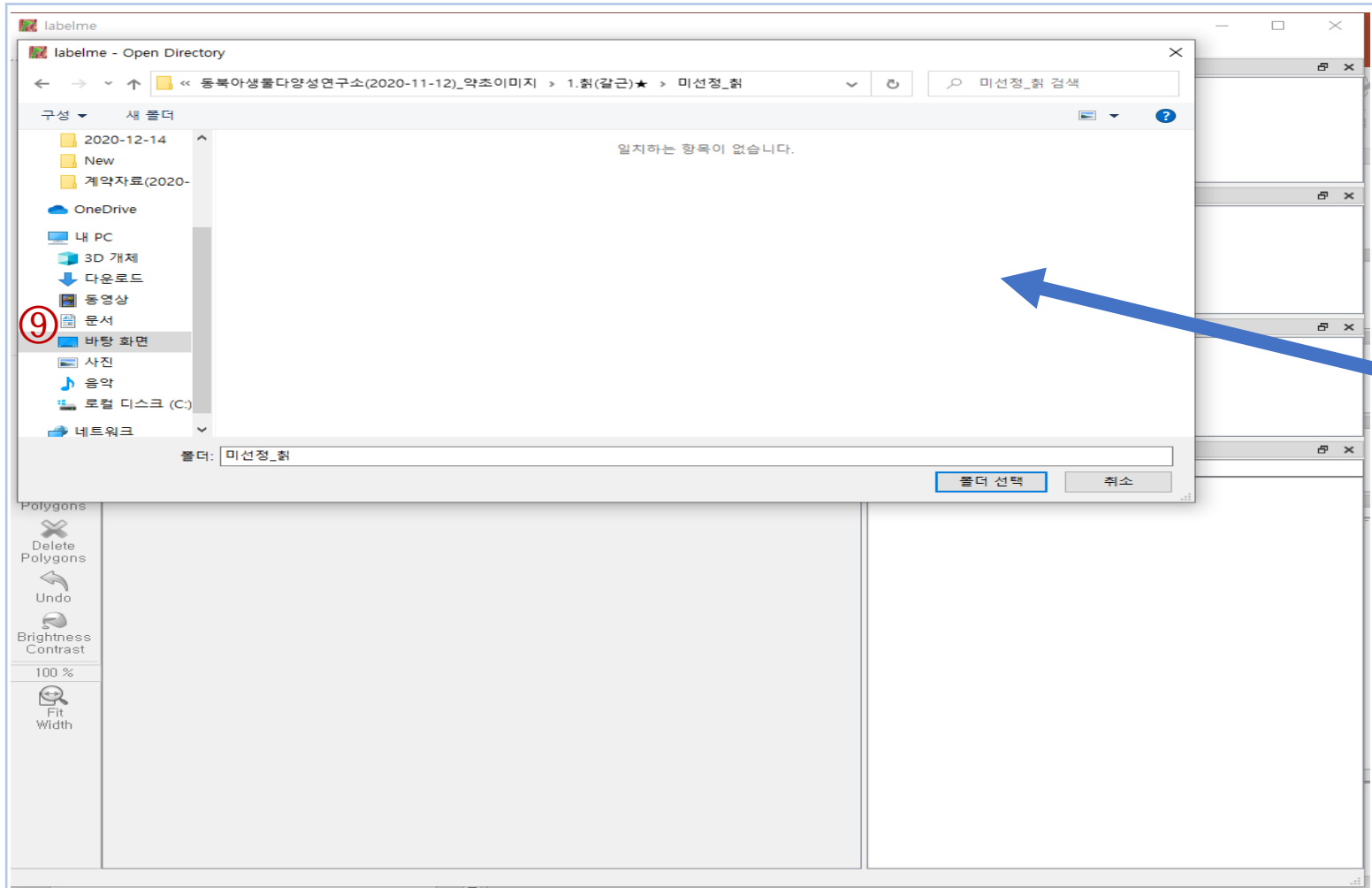
#### 가공 작업 6

- ⑧ 앞장에 File에서  
Save with Image Data 를 반드시 클릭  
Save Automatically을 클릭한 후  
**Click the Open Dir 선택**  
하면 다운로드된 파일을 불러옴.



### 3. 데이터 가공 작업 [Labelme]

#### 가공 작업 7

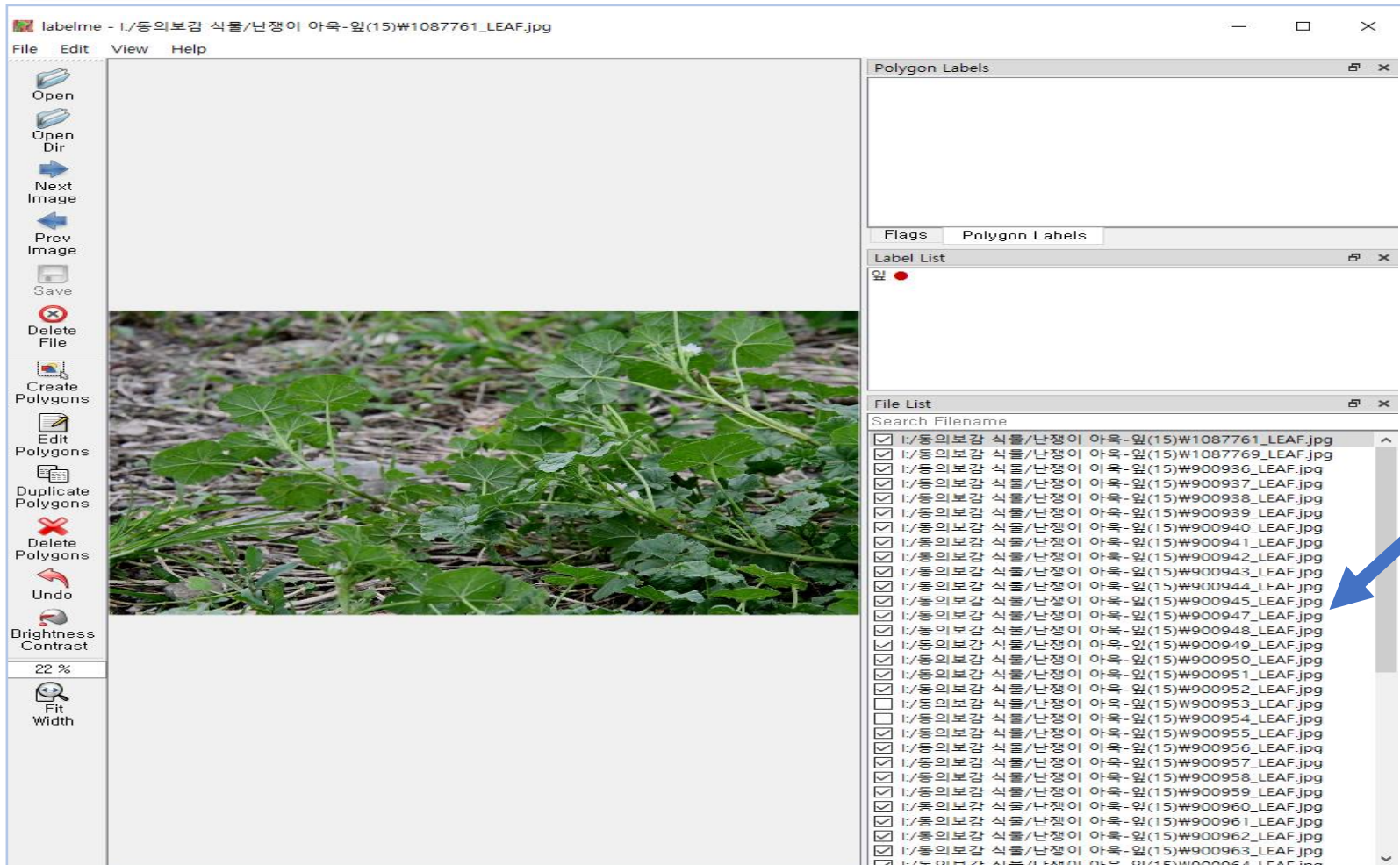


#### 가공 작업 7

⑨ 이미지 디렉토리 및 폴더 버튼 클릭  
다음 가공 대상 사진으로 쉽게 이동

### 3. 데이터 가공 작업 [Labelme]

#### 가공 작업 8

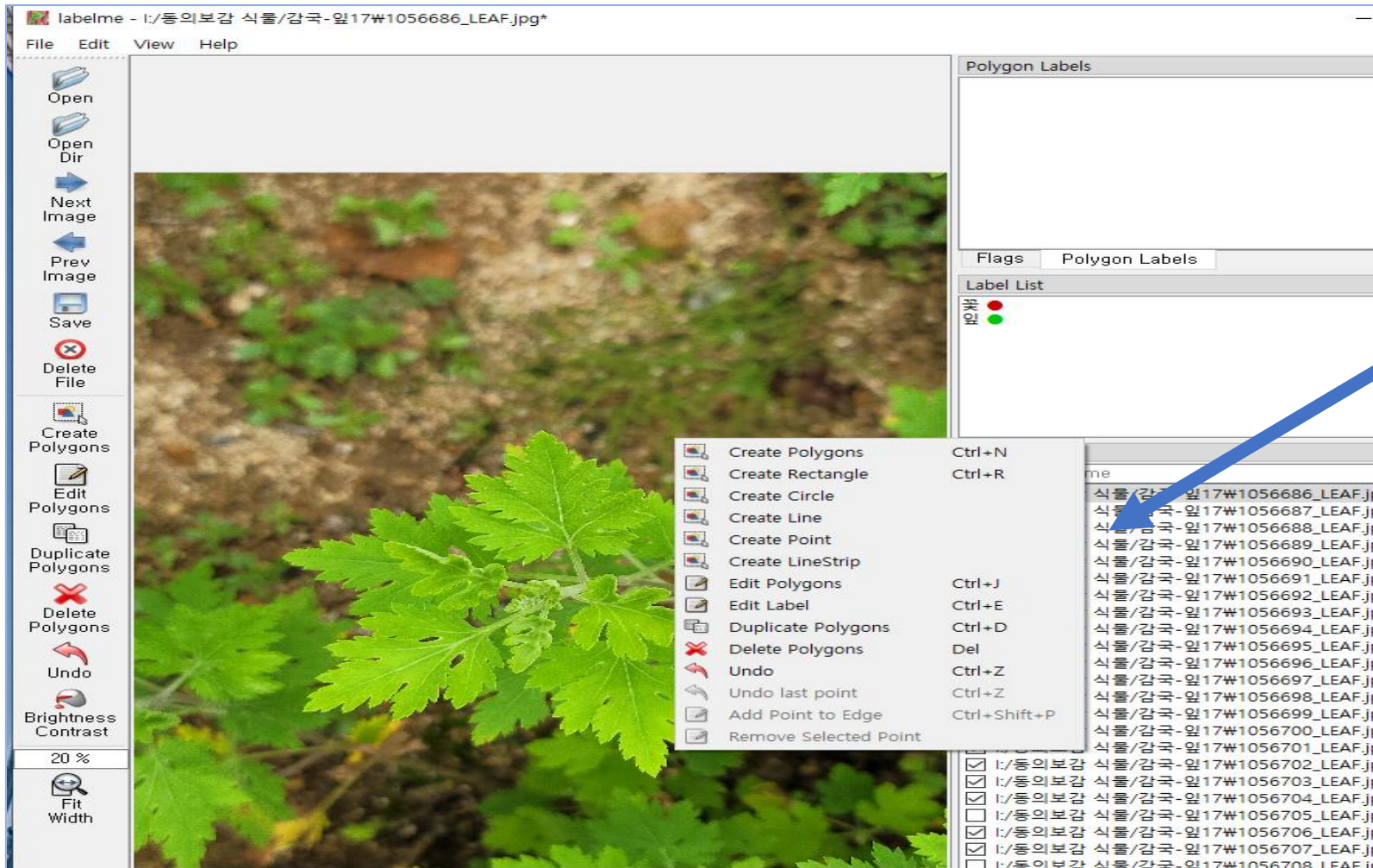


#### 가공 작업 8

⑩ 앞의 모든 이미지가 표시됩니다

### 3. 데이터 가공 작업 [Labelme]

#### 가공작업 9



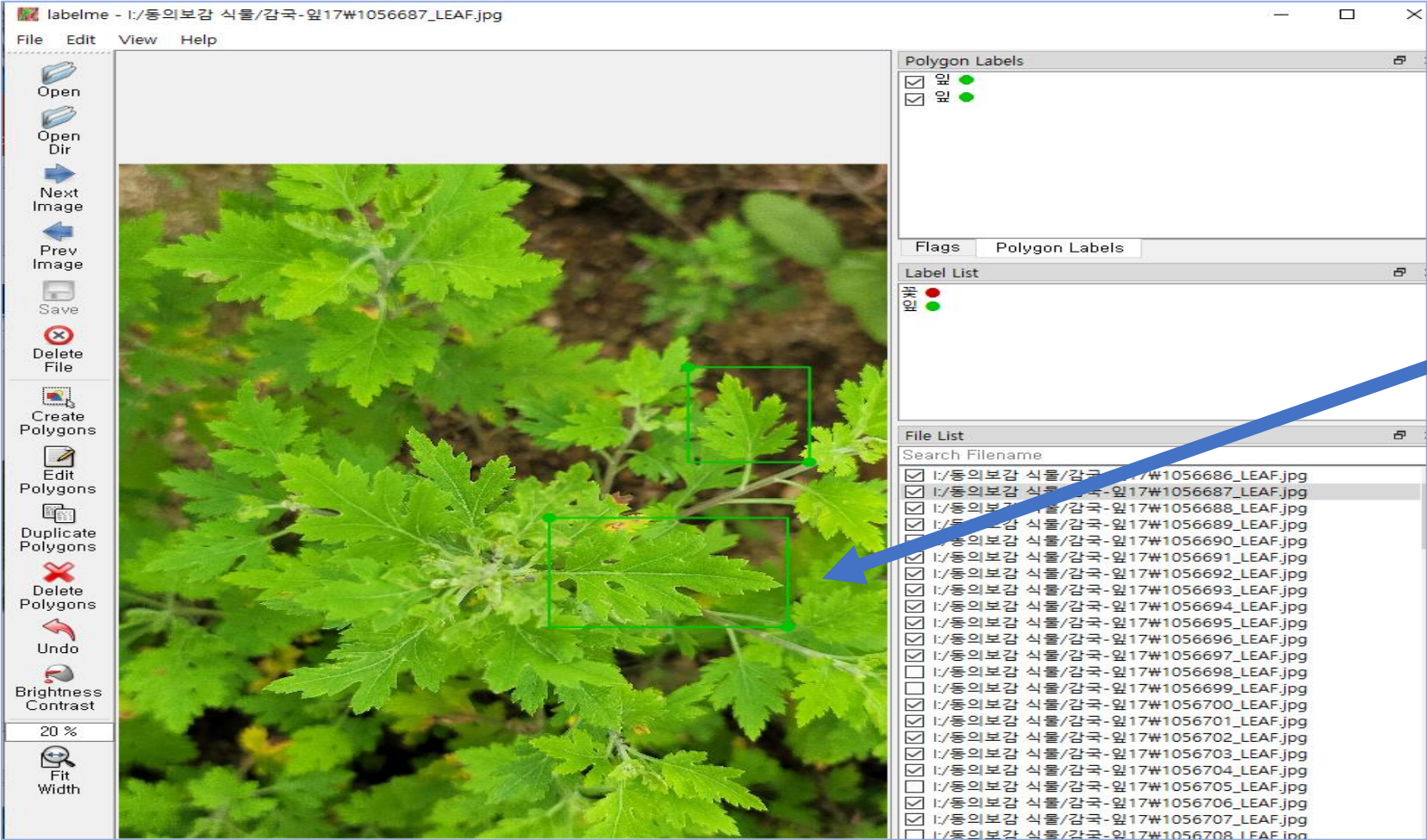
#### 가공 작업 9

- ⑪ 다운받은 이미지를 불러들인 다음  
마우스 오른쪽 버튼 클릭,  
그리고 사각형 만들기를 클릭



# 3. 데이터 가공 작업 [Labelme]

## 가공작업 10



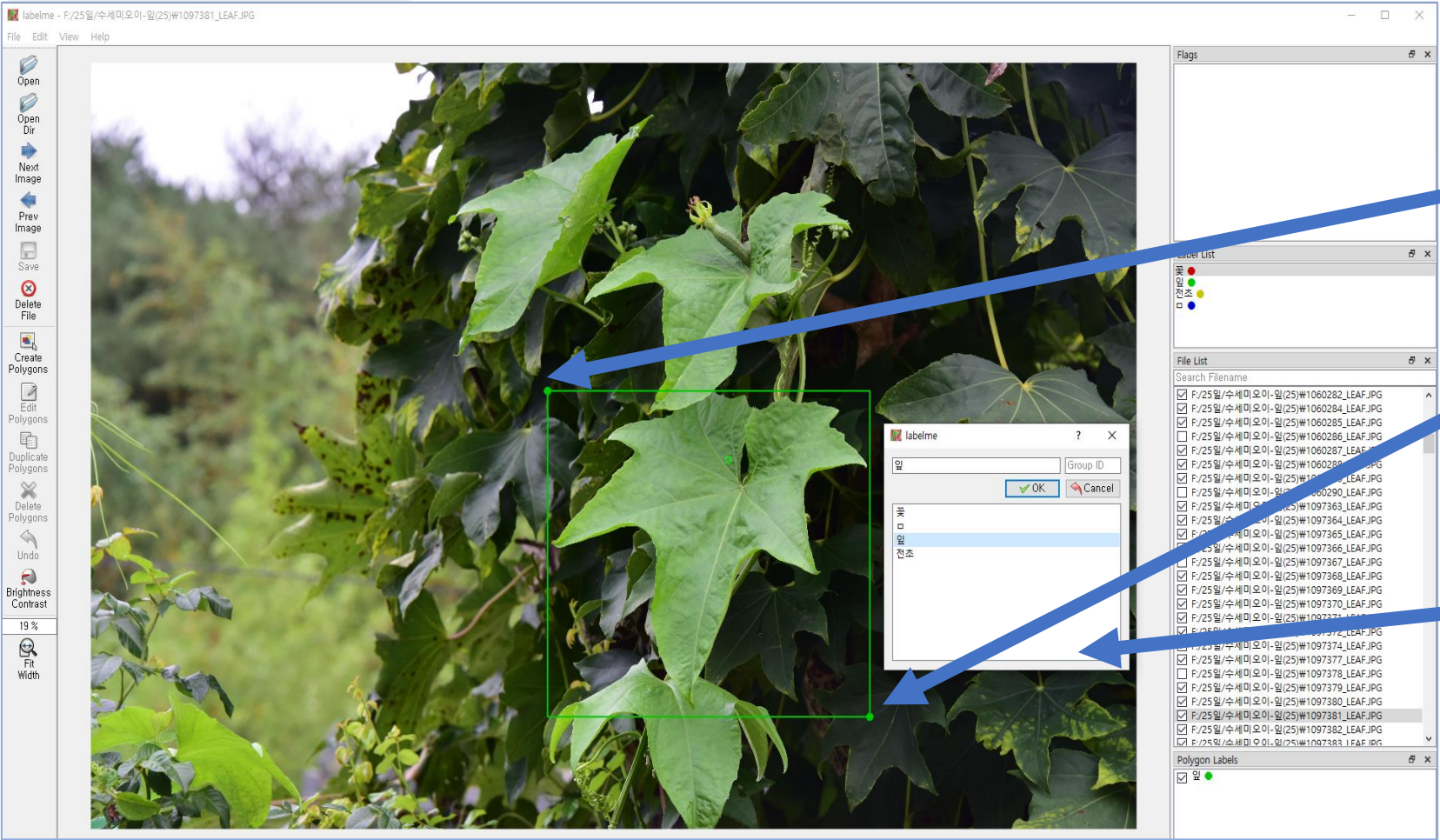
가공 작업 10

⑫ 사각형이 만들어진  
이미지를 확인



### 3. 데이터 가공 작업 [Labelme]

#### 가공작업 11



#### 가공 작업 11

⑬ 마우스 버튼을 클릭

⑭ 그리기 및 클릭  
다시 마우스 버튼

⑮ LabelMe 상자가 표시됨

### 3. 데이터 가공 작업 [Labelme]

#### 가공작업 12



가공 작업 12

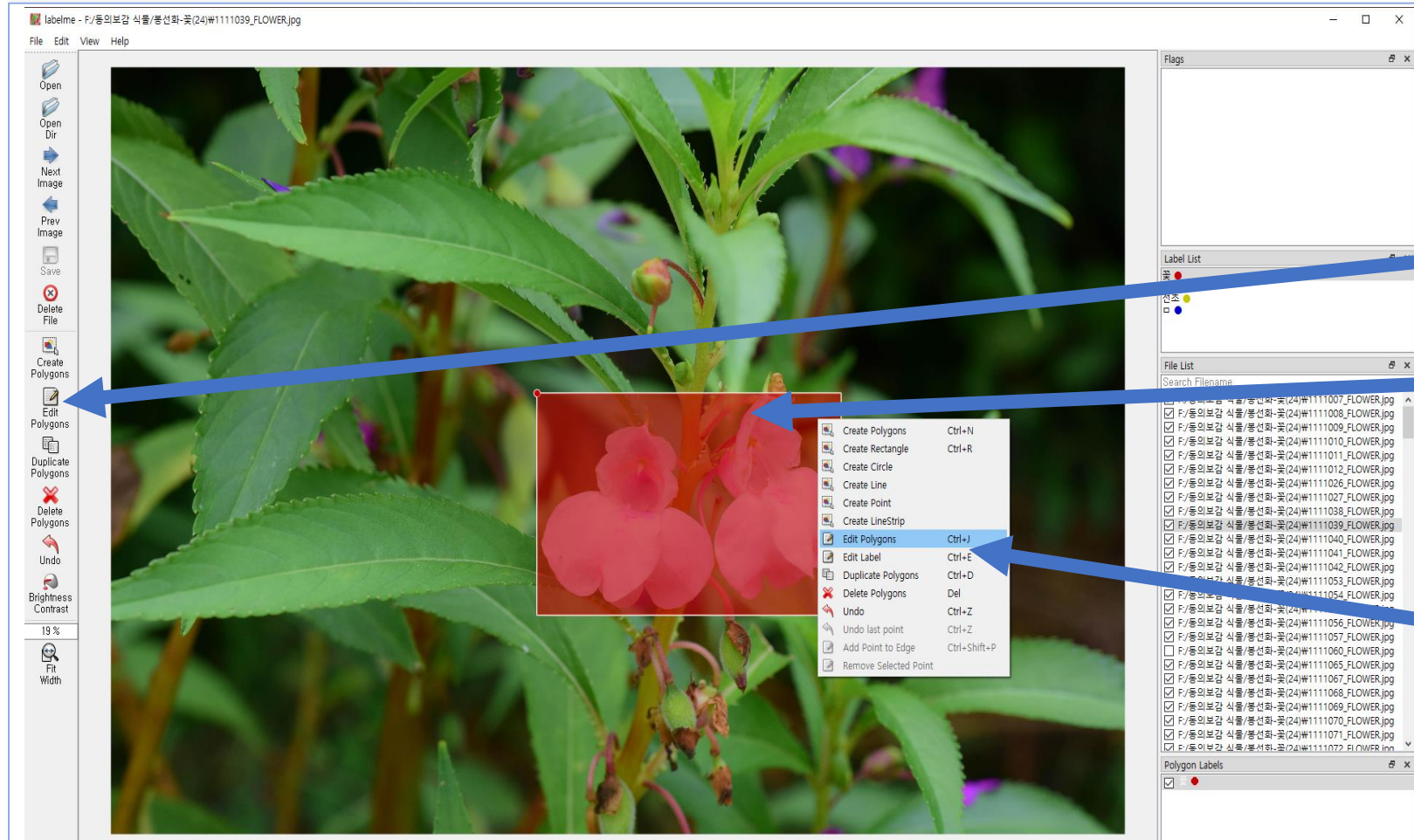
⑩ 이름 쓰기 : \_잎

⑪ OK (확인버튼) 클릭



## 4. 결과물 [1차 가공]

### LabelMe Guide



수정 또는 이동 또는 삭제

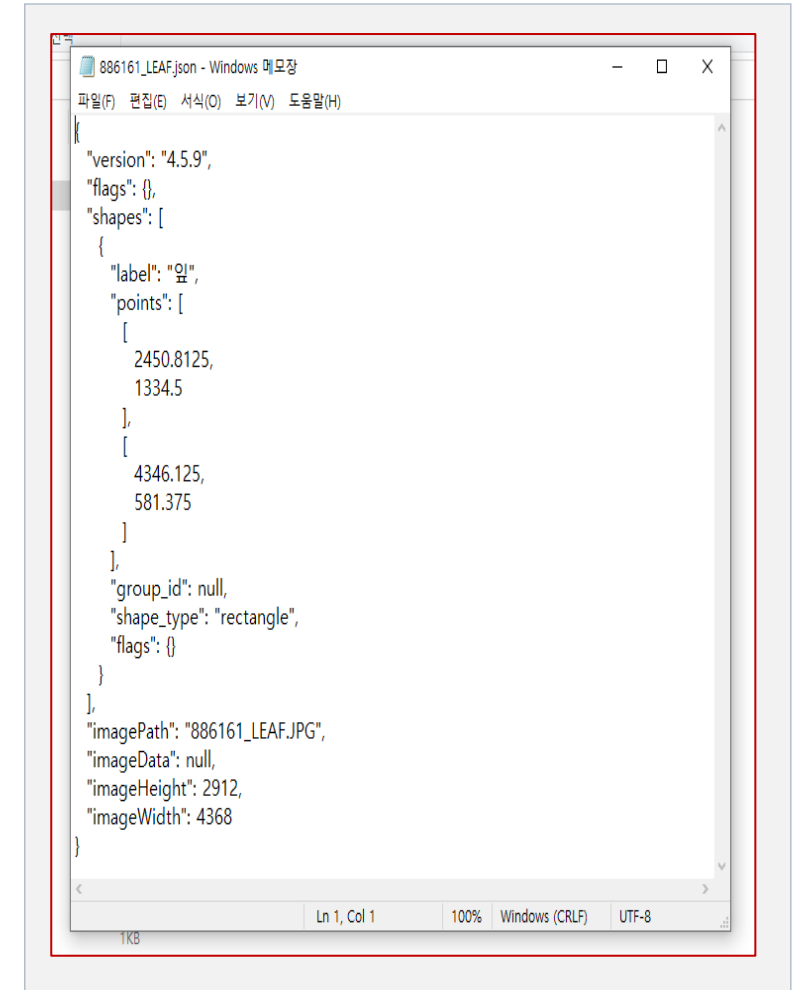
클릭 'Edit Polygons' icon

마우스 커서를 상자로 이동하고,  
손 모양에 따라 커서 변경, 왼쪽 버튼을  
클릭하고 이동

상자에서 오른쪽 버튼을 클릭하고,  
기능을 선택하고, 그리고 수정

## 4. 결과물 [1차 가공]

### 결과물





## 4. 결과물 [1차 가공]

꽃 labeling 예시

→ **반려**되는 예시



➡ 바운딩 박스 안에 여러개체를 **create Rectangles** 할 수 있음. (단축키: Ctrl+R[박스])



➡ 바운딩 박스 안에 같은 종류의 꽃이 아닌 다른종류의 꽃이 포함되어 있거나 다른 이물질이 있는경우 **반려**됨.



## 4. 결과물 [1차 가공]

### 꽃 labeling 예시



바운딩 박스 안에 여러개체를 **create Rectangles** 할 수 있음.



꽃잎의 끝부분이 잘려 있지만 원본의 90%이상이 유지하고 있어 **create Rectangles** 가능함.

## 4. 결과물 [1차 가공]

### 꽃 labeling 예시



바운딩 박스 안에 여러개체를 **create Rectangles**  
할 수 있음. ( 단축키: Ctrl+R)



바운딩 박스 안에 여러개체를 **create polygons**  
할 수 있음. ( 단축키: Ctrl+N)



## 4. 결과물 [1차 가공]

### 앞 labeling 예시



바운딩 박스 안에 여러개체에 앞을 **create polygons**  
할 수 있음. ( 단축키: Ctrl+N)



바운딩 박스 안에 앞을 **create Rectangles**  
할 수 있음. ( 단축키: Ctrl+R)



## 4. 결과물 [1차 가공]

### 앞 labeling 예시



바운딩 박스 안에 여러개체에 앞을 **create Rectangles**  
할 수 있음. ( 단축키: Ctrl+R)



바운딩 박스 안에 여러개체에 앞을 **create polygons**  
할 수 있음. ( 단축키: Ctrl+N)



## 4. 결과물 [1차 가공]

### 앞 labeling 예시



바운딩 박스 안에 여러개체에 앞을 **create Rectangles**  
할 수 있음. ( 단축키: Ctrl+R)



바운딩 박스 안에 여러개체에 앞을 **create Rectangles**  
할 수 있음. ( 단축키: Ctrl+R)



## 4. 결과물 [1차 가공]

### 앞 labeling 예시



바운딩 박스 안에 여러개체에 앞을 **create Rectangles** 할 수 있음. ( 단축키: Ctrl+R)



바운딩 박스 안에 여러개체에 앞을 **create Rectangles** 할 수 있음. ( 단축키: Ctrl+R)



## 4. 결과물 [1차 가공]

### 열매 labeling 예시



➡ 바운딩 박스 안에 여러개체에 열매를 **create Rectangles** 할 수 있음. ( 단축키: Ctrl+R)



➡ 바운딩 박스 안에 여러개체에 열매를 **create Rectangles** 할 수 있음. ( 단축키: Ctrl+R)



## 4. 결과물 [1차 가공]

### 열매 labeling 예시



➡ 바운딩 박스 안에 여러개체에 열매를 **create polygons** 할 수 있음. ( 단축키: Ctrl+N)



➡ 바운딩 박스 안에 여러개체에 열매를 **create polygons** 할 수 있음. ( 단축키: Ctrl+N)



## 4. 결과물 [1차 가공]

### 전초 labeling 예시



➡ 바운딩 박스 안에 꽃 or 잎 or 열매를 최소한의 여백을 제거 후 **create polygons** 할 수 있음. ( 단축키: Ctrl+N)



➡ 바운딩 박스 안에 꽃 or 잎 등을 **create Rectangles** 할 수 있음. ( 단축키: Ctrl+R)



## 4. 결과물 [1차 가공]

### 전초 labeling 예시



바운딩 박스 안에 꽃, 잎을 **create polygons**  
할 수 있음. ( 단축키: Ctrl+N)



바운딩 박스 안에 여러개체에 열매,잎을 **create Rectangles**  
할 수 있음. ( 단축키: Ctrl+R)



## 4. 결과물 [1차 가공]

### 전초 labeling 예시



바운딩 박스 안에 꽃, 잎을 **create Rectangles**  
할 수 있음. ( 단축키: Ctrl+R)

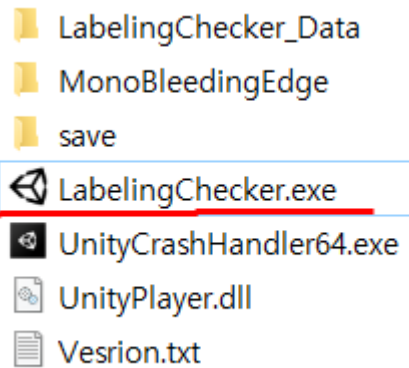


바운딩 박스 안에 꽃, 잎을 **create Rectangles**  
할 수 있음. ( 단축키: Ctrl+R)

## 4. 라벨링 검증 툴

### 라벨링 검증 툴 사용 예시

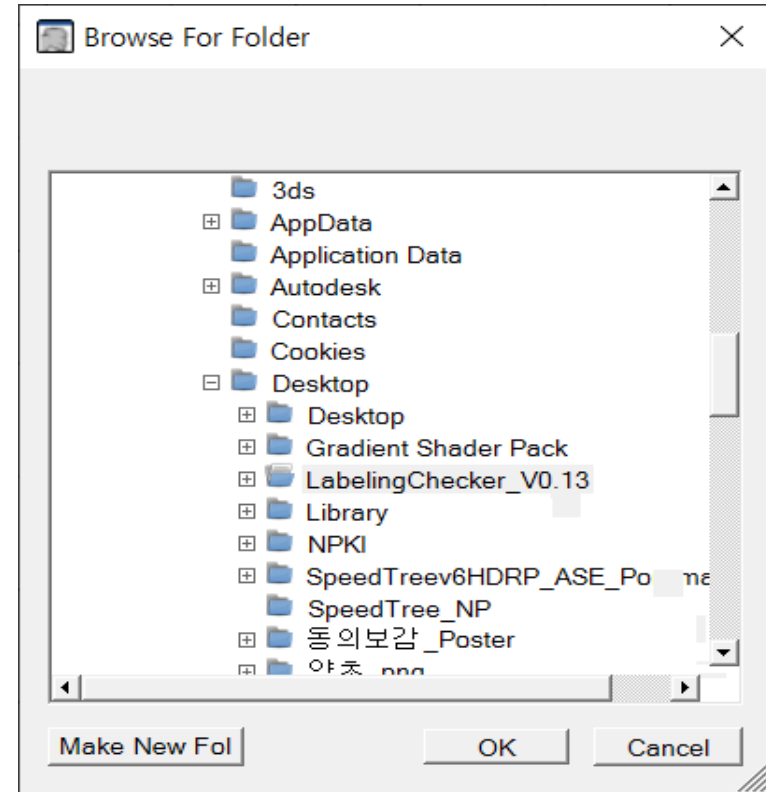
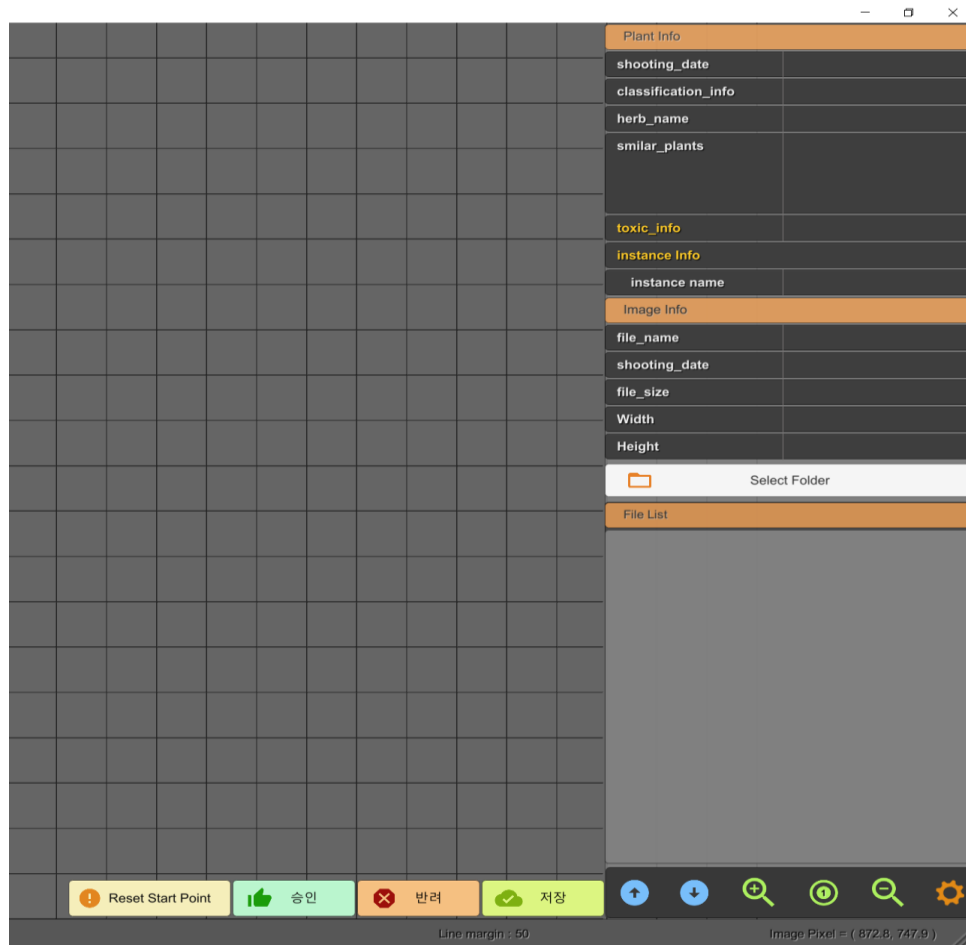
1. 압축파일을 풀고 LabelingChecker.exe 실행



## 4. 라벨링 검증 툴

### 라벨링 검증 툴 사용 예시

2. Select Folder 버튼클릭 후 폴더 선택

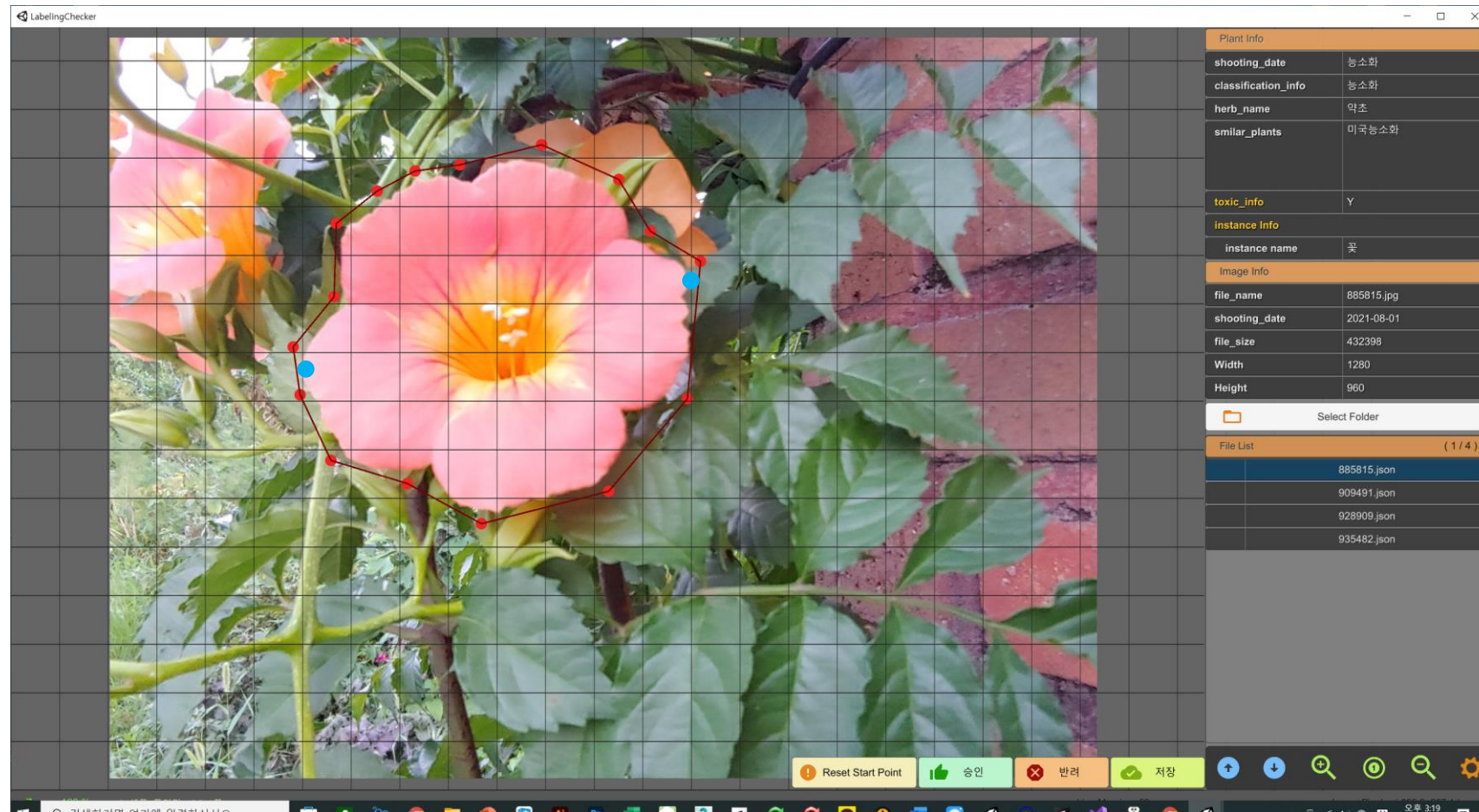




## 4. 라벨링 검증 툴

### 라벨링 검증 툴 사용 예시

3. 마우스 오른쪽 버튼으로 이미지 객체의 가로나 세로의 가장 튀어나온 부분에 ●을 찍는다

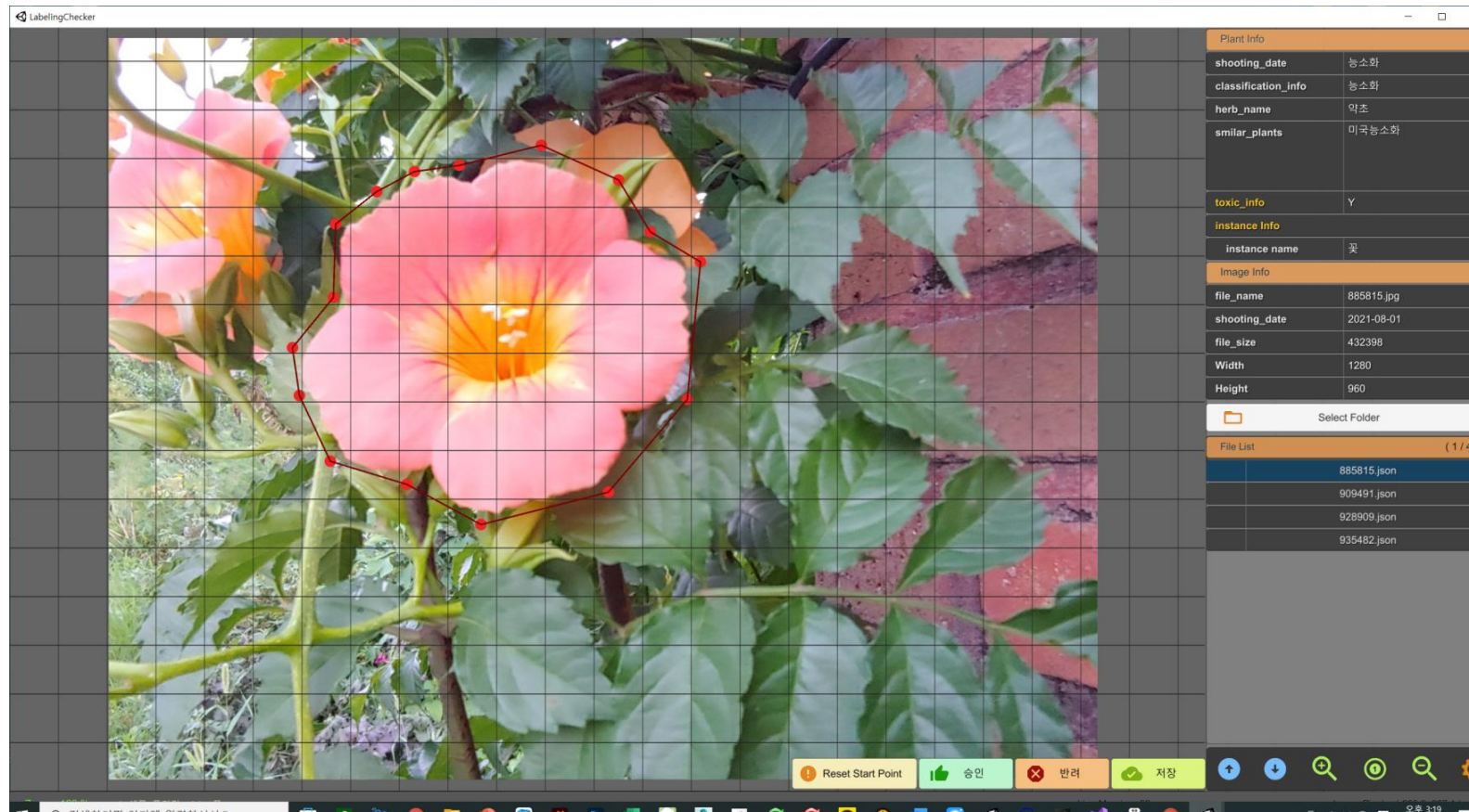


## 4. 라벨링 검증 툴

### 라벨링 검증 툴 사용 예시

- 을 찍고 나면 잠시후 그리드가 새롭게 생성됨

오브젝트 객체와 라벨링의 간격이 1 그리드 안에 들어 오면 합격 처리됨(오브젝트의 면적이 라벨링 면적의 80%로 자동 계산됨 )





# 감사합니다

