



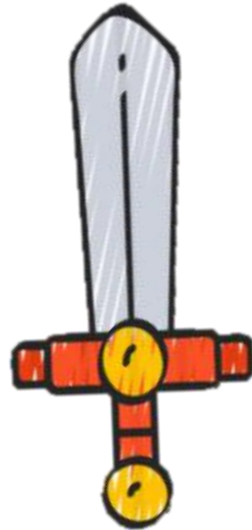
인턴십 최종발표(6/24) **이미지 카빙 자동화 프로그램**

융합보안공학과 7팀 성연수, 신나연, 한아름

INDEX

- 01 프로젝트작품 소개
- 02 프로젝트작품 구현
- 03 프로젝트작품 시연
- 04 보완점
- 05 역할분담

01 프로젝트작품 소개 | 제안 배경 (필요성 – 문제관점)



01 프로젝트작품 소개 | 제안 배경 (필요성 – 요구관점)

디지털 포렌식 시대, 이미지·동영상 자동 분석 수요 ↑

이유지 | 2017.05.23

공유 3 | 댓글 0

언어 선택 ▼ Google 번역에서 제공

가+ 가-

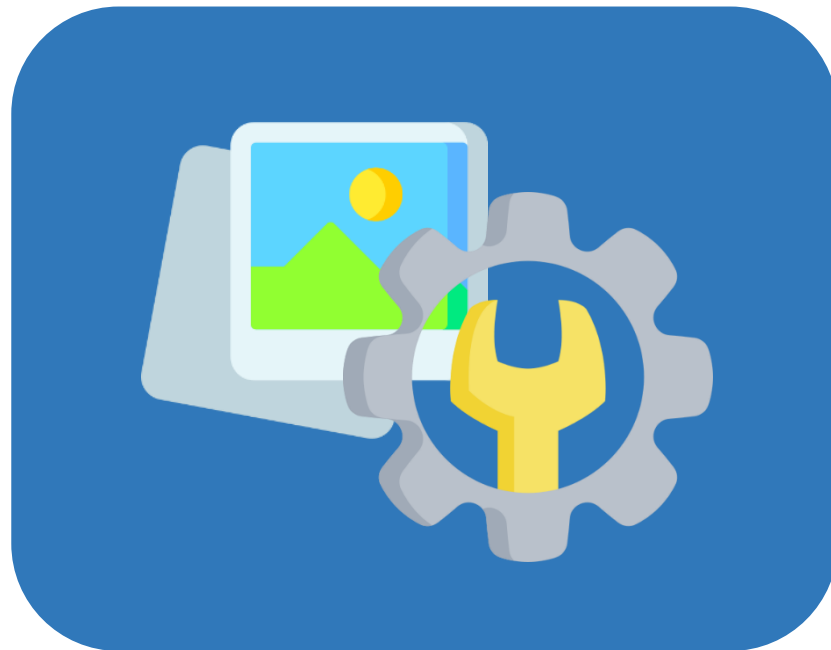
스마트폰과 디지털카메라, 전국 곳곳에 설치돼 있는 CCTV·영상감시 카메라를 통해 생성되고, 컴퓨터와 인터넷상에 저장·게재·배포되는 디지털 이미지와 동영상이 급증하고 있다.

이에 따라 수사기관에서도 사건 해결을 위한 단서나 범죄 증거를 찾기 위해 분석해야 하는 미디어의 양이 기하급수적으로 늘어나는 추세다.

범죄 증거를 찾기 위해
분석해야 할 **이미지의 증가**



디지털포렌식과 **자동화 tool**에 대한
요구 증가



이미지카빙 + 자동분석 및 정리



이미지 카빙 자동화 프로그램

01 프로젝트작품 소개 | 편리성, 필요성, 유용성



이미지 포맷 별 카빙 설계
(JPG, PNG, BMP등)



카빙 후 CSV를 통한
자동화 정리



시간정보(생성, 수정, 삭제 시간),
파일 크기, 이미지 등 자동 정리

01 프로젝트작품 소개 | 편리성, 필요성, 유용성

편리성

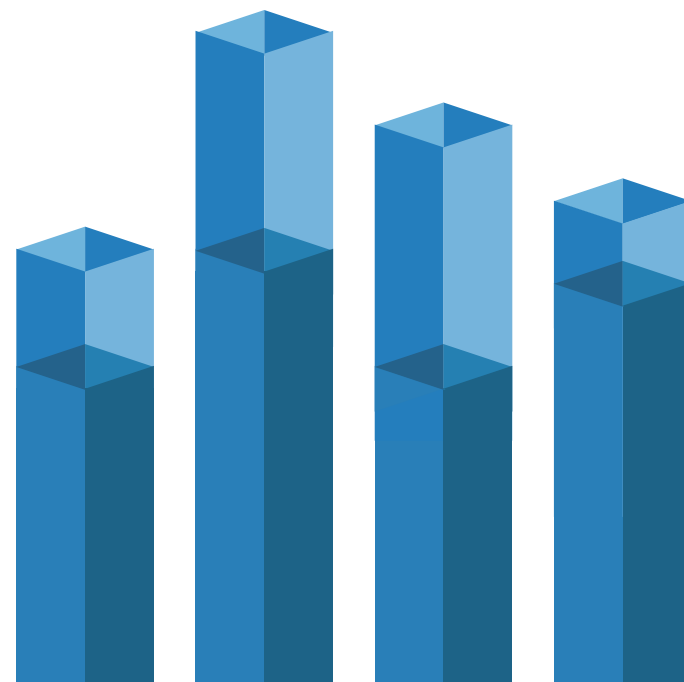
UI가 간단하기 때문에 쉽게 이미지 카빙을 할 수 있음
삭제된 파일을 복구하고 싶을 때 쉽게 사용 가능

필요성

이미지 속성정보를 분석해 자동화된 이미지 카빙 도구를 만들어
프로그램 언어를 모르는 사람들도 쉽게 사용 가능

유용성

삭제한 파일(JPEG, PNG, BMP)을 이미지 카빙을 통해 복구 가능
이미지 속성정보를 CSV 파일로 추가 제공해
카빙 완료된 파일들의 속성정보를 비교하며 한번에 볼 수 있음



1 이미지 카빙

JPEG, PNG, BMP 이미지 파일을 파일 구조에 따라 두 가지 방법으로 카빙

2 이미지 속성 정보 제공

카빙 완료 후 속성정보를 추출하여 이미지 속성 정보를 제공

3 CSV 파일 제공

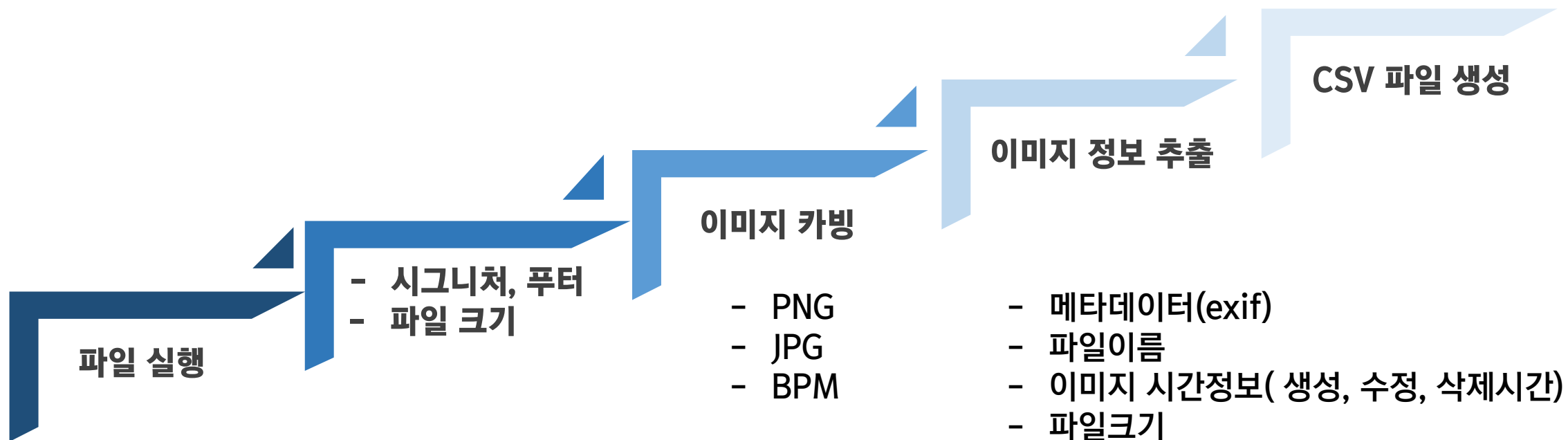
카빙 결과를 CSV 파일로 확인 가능

4 카빙 Tool 제작(exe 실행 파일)

사용하기 쉬운 GUI 형태의 실행파일을 제공

02 프로젝트작품 구현 | 제품의 동작 과정

제품의 동작 과정



프로그램 사용환경

사용 환경

운영체제

Windows 10

파이썬 버전

Python 3.7

에디터

PyCharm

사용 라이브러리

이미지 카빙 라이브러리

imghdr

이미지 유형 판단

binascii

바이너리-ASCII 간의 변환

이미지 정보 추출 라이브러리

PIL

Python Imaging Library, ExifTags

Pillow

PIL 후속 버전

실행파일 생성 라이브러리

PyInstaller

실행 파일 생성

PyQt5

GUI 프로그램

엑셀 CSV 생성 라이브러리

CSV

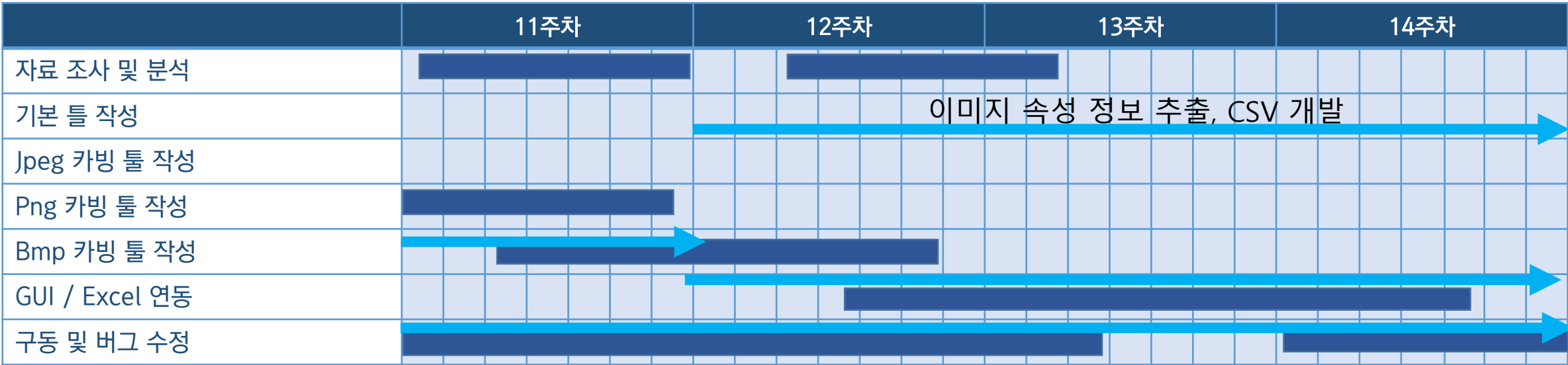
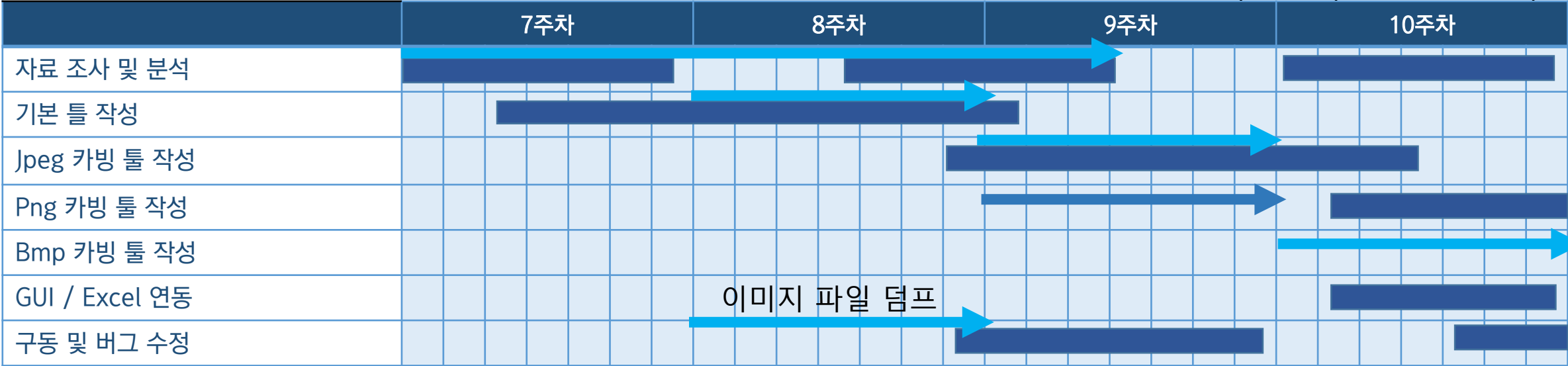
csv 파일 생성

Pandas

csv 파일 생성

02 프로젝트작품 구현 | 프로젝트 계획과 달성

→ : 실제 프로젝트 활동 진행 사항



*일정은 상황에 따라 변동 가능성 있음

02 프로젝트작품 구현 | 최종 결과물 코드

최종 결과물 코드(github\sungshin2020\imgCarving\SUB MASTER\gui)

sys_mainwindow.ui

File Carving 프로그램의 UI 파일

sys_mainwindow.py

File Carving 프로그램

ImageCarving_for_jpg_png.py

헤더 푸터 방식으로 카빙하는 JPG, PNG 카빙 알고리즘

ImageCarving_for_bmp.py

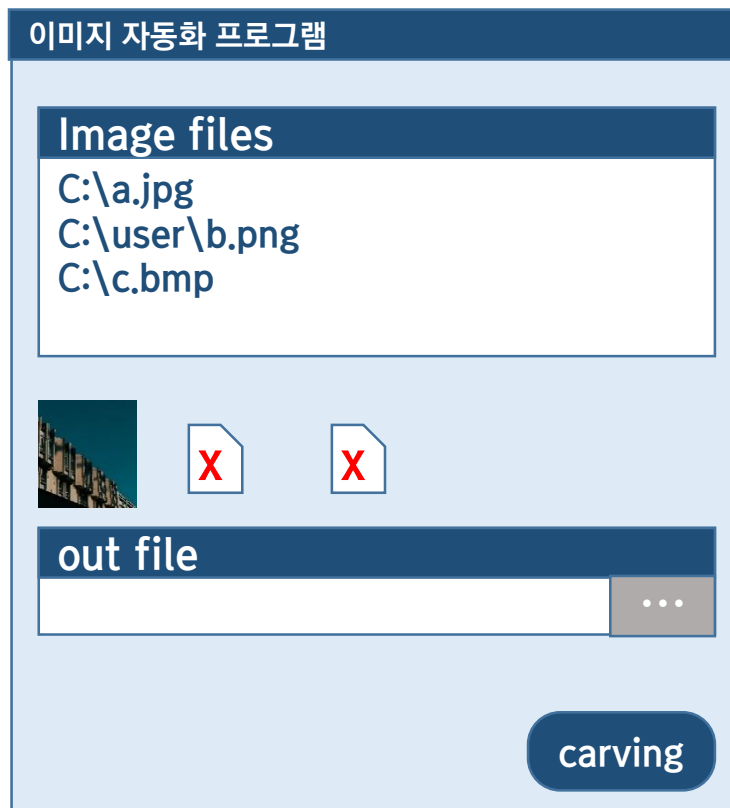
파일 크기 방식으로 카빙하는 BMP 카빙 알고리즘

Image_EXIF.py

파일 속성 정보 추출과 CSV 파일 생성 코드

02 프로젝트작품 구현 | 제안서에서의 기능과 최종 발표에서의 기능 차이

이미지 카빙 자동화 프로그램(GUI) | 예상 결과물



*화면 구성이나 옵션 등은 변동가능성 높음

02 프로젝트작품 구현 | 제안서에서의 기능과 최종 발표에서의 기능 차이

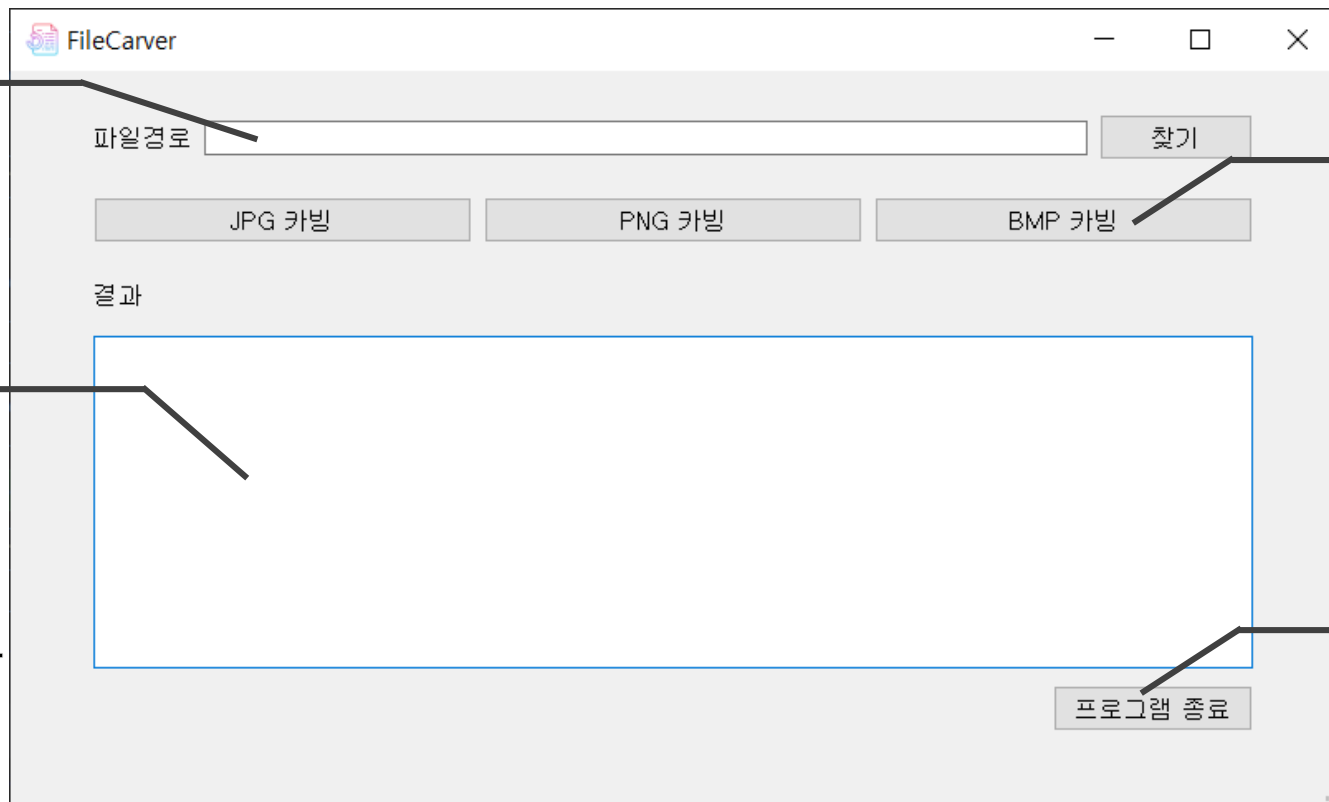
이미지 카빙 자동화 프로그램(GUI) | 구현 완료

파일탐색기

이미징한 파일을 불러옴

실행 화면

파일 탐색 결과,
파일 카빙 결과,
CSV 생성 결과를
사용자가 확인할 수 있음



확장자 별 파일 카빙

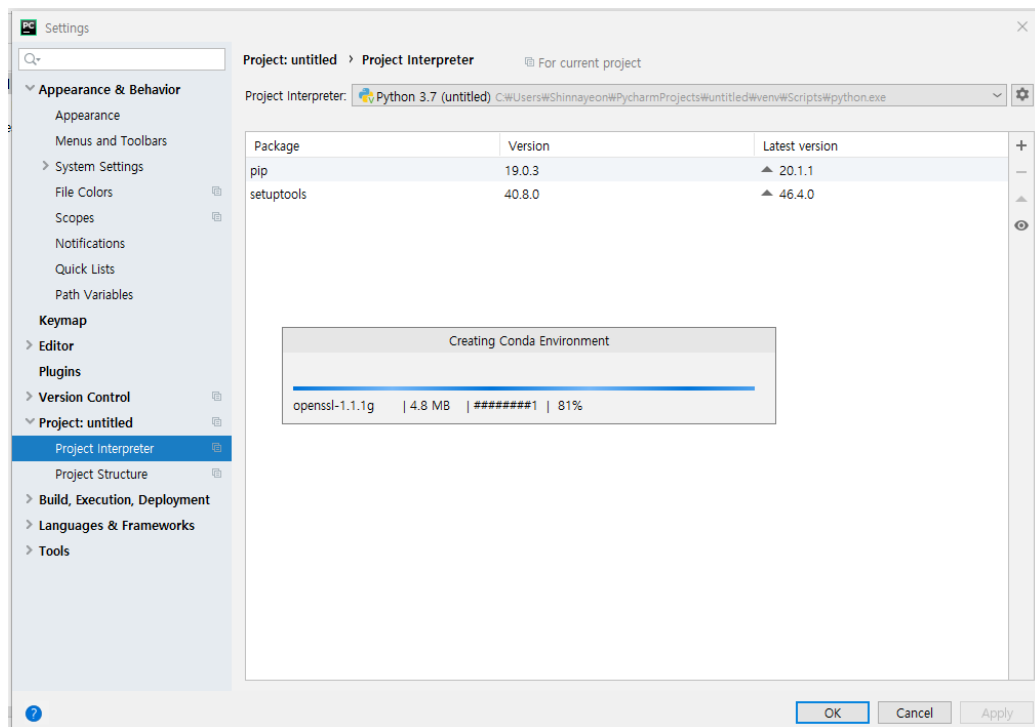
원하는 확장자를 클릭하면
파일 카빙이 시작됨

나가기

프로그램을 종료함

02 프로젝트작품 구현 | 제안서에서의 기능과 최종 발표에서의 기능 차이

카빙 알고리즘 제작 | 문제점



Anaconda를 이용하여
pyQt 설치 중에 오류 발생

```
C:\Users\Shinnayeon\AppData\Local\Programs\Python\Python37\Scripts>pip install PyQt5
Collecting PyQt5
  Downloading https://files.pythonhosted.org/packages/d7/8e/5fa1dd8095728fa754e96633d4c97e0283fb0be5ab3a0a25f7df054deff1/PyQt5-5.14.2-5.14.2-cp35-cp36-cp37-cp38-none-win_amd64.whl (52.9MB)
    | 49.6MB 6.4MB/s eta 0:00:01
```

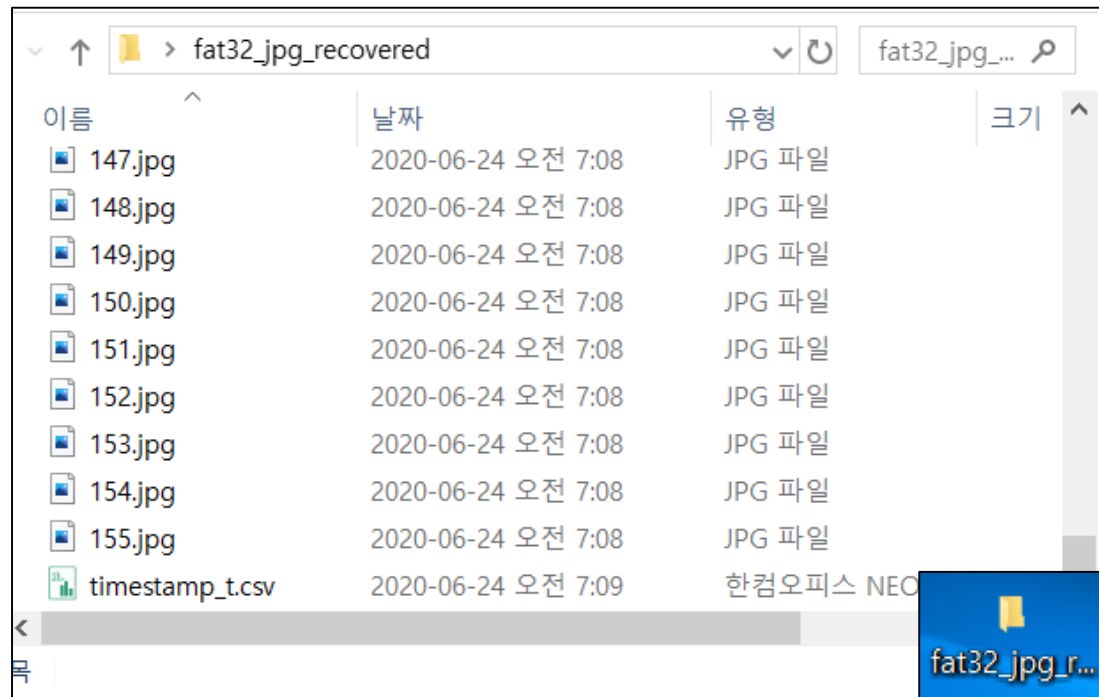
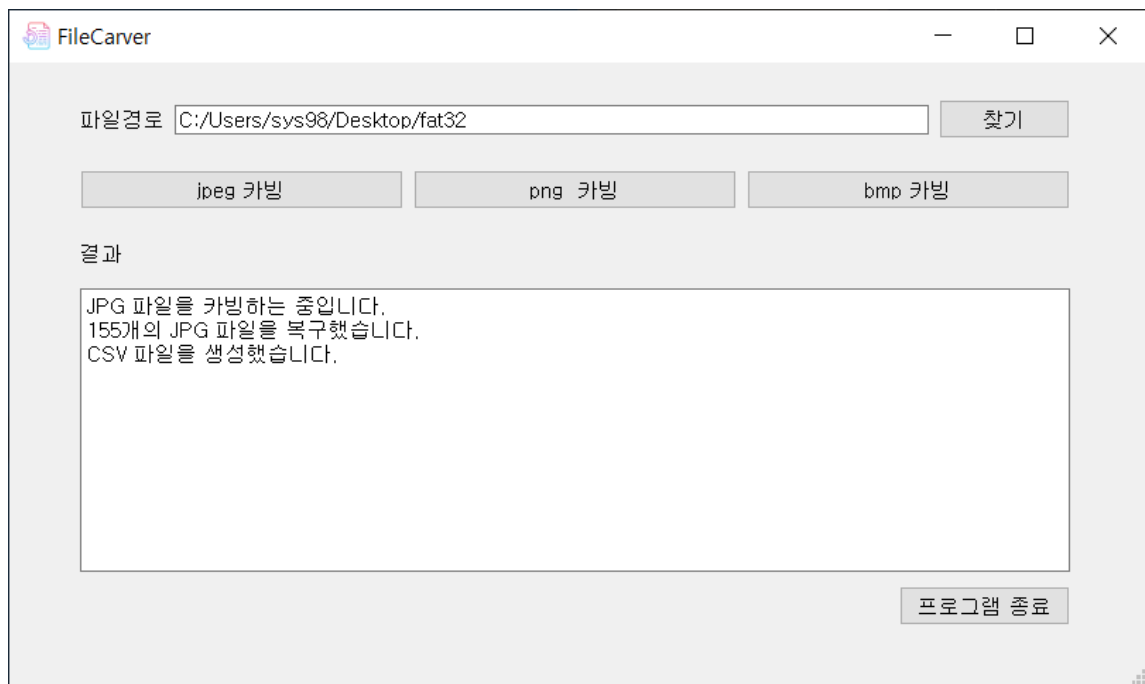
```
C:\Users\Shinnayeon\AppData\Local\Programs\Python\Python37\Scripts>pip install PyQt5
Collecting PyQt5
  Downloading https://files.pythonhosted.org/packages/d7/8e/5fa1dd8095728fa754e96633d4c97e0283fb0be5ab3a0a25f7df054deff1/PyQt5-5.14.2-5.14.2-cp35-cp36-cp37-cp38-none-win_amd64.whl (52.9MB)
    | 52.9MB 123kB/s
Collecting PyQt5-sip<13,>=12.7 (from PyQt5)
  Downloading https://files.pythonhosted.org/packages/11/9f/093f7aa50af963a6cc825d1392770ea4ad821f175de1cd8bcb6646be27a6/PyQt5_sip-12.7.2-cp37-cp37m-win_amd64.whl (58kB)
    | 61kB 3.8MB/s
Installing collected packages: PyQt5-sip, PyQt5
  WARNING: The scripts pyupdate5.exe, pyrcc5.exe and pyuic5.exe are installed in 'c:\Users\shinnayeon\AppData\Local\Programs\Python\Python37\Scripts' which is not on PATH.
  Consider adding this directory to PATH or, if you prefer to suppress this warning, use --no-warn-script-location.
Successfully installed PyQt5-5.14.2 PyQt5-sip-12.7.2
WARNING: You are using pip version 19.2.3, however version 20.1.1 is available.
You should consider upgrading via the 'python -m pip install --upgrade pip' command.
C:\Users\Shinnayeon\AppData\Local\Programs\Python\Python37\Scripts>
```

```
C:\Users\Shinnayeon\AppData\Local\Programs\Python\Python37\Scripts>pip install --upgrade pip
Collecting pip
  Downloading https://files.pythonhosted.org/packages/43/84/23ed5a1796480a6f1a2d38f2802901d078266bda38388954d01d3f2e821c/pip-20.1.1-py2.py3-none-any.whl (1.5MB)
    | 1.5MB 656kB/s
Installing collected packages: pip
  Found existing installation: pip 19.2.3
  Uninstalling pip-19.2.3:
    Successfully uninstalled pip-19.2.3
```

CMD를 통해 파이썬 경로에 직접 설치

02 프로젝트작품 구현 | 제안서에서의 기능과 최종 발표에서의 기능 차이

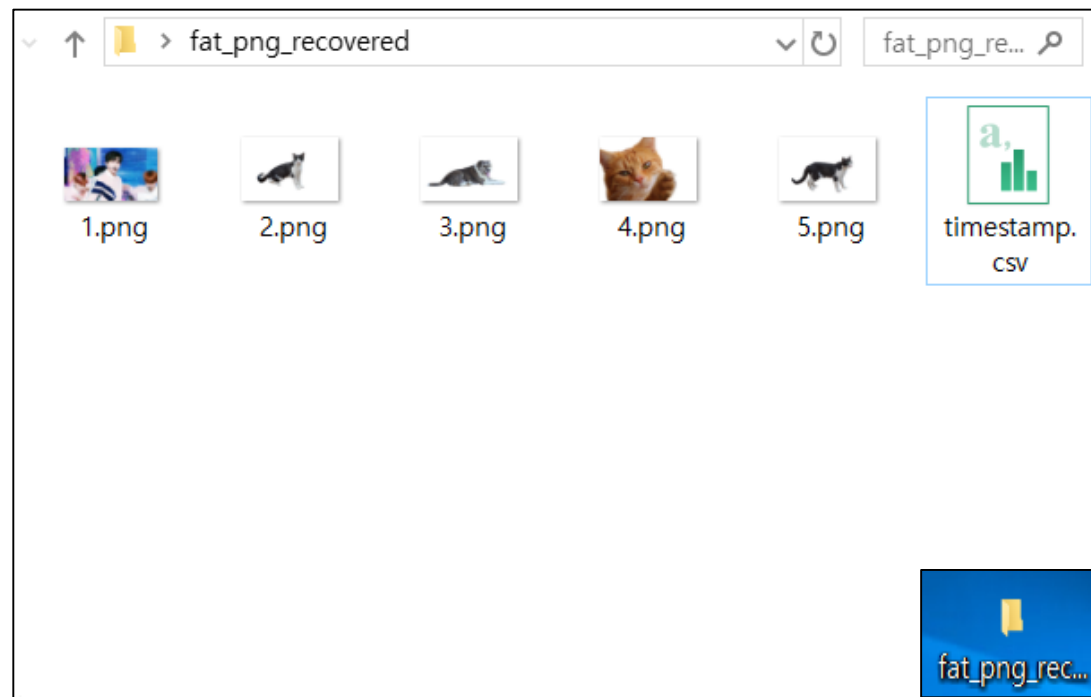
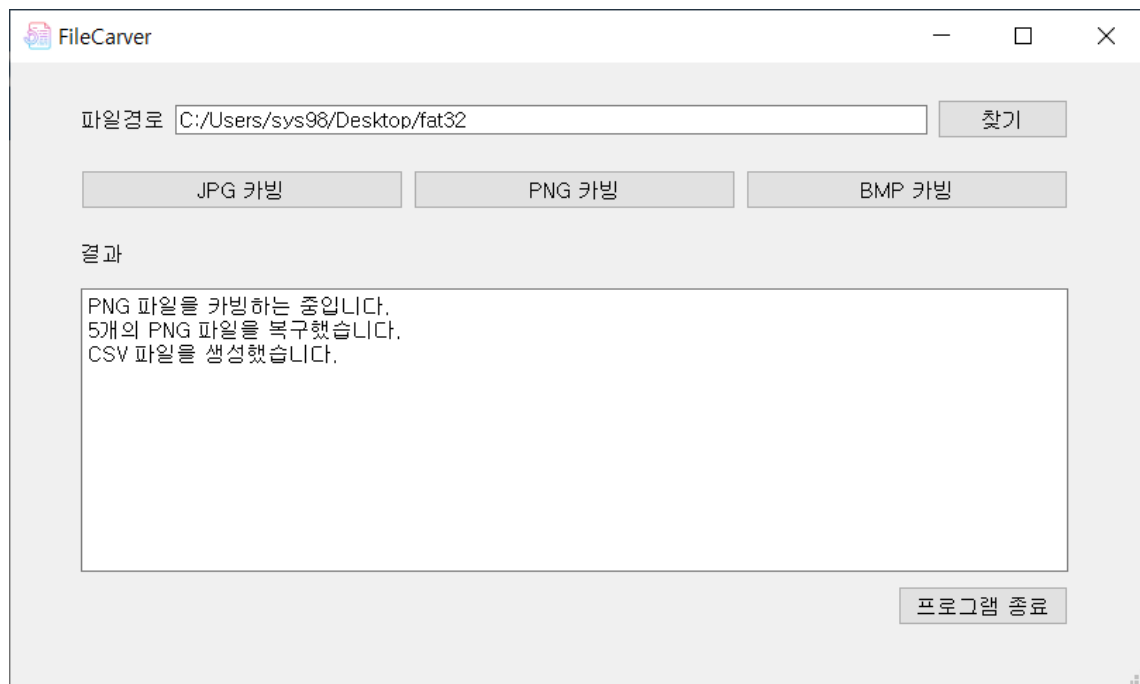
JPG 알고리즘 | 구현 완료



1. [파일 경로]에서 FAT32 파일시스템을 가지는 덤프 파일[fat32]를 선택한다.
2. [JPG 카빙] 버튼을 클릭하면 JPG 파일을 카빙한다.
3. 결과 화면을 통해 155개의 JPG 파일을 복구했음을 알 수 있다.
4. 1에서 선택한 [파일경로]에 [덤프 파일 이름_jpg_recovered]라는 폴더가 생기고, 이 폴더에서 이미지를 확인할 수 있다.

02 프로젝트작품 구현 | 제안서에서의 기능과 최종 발표에서의 기능 차이

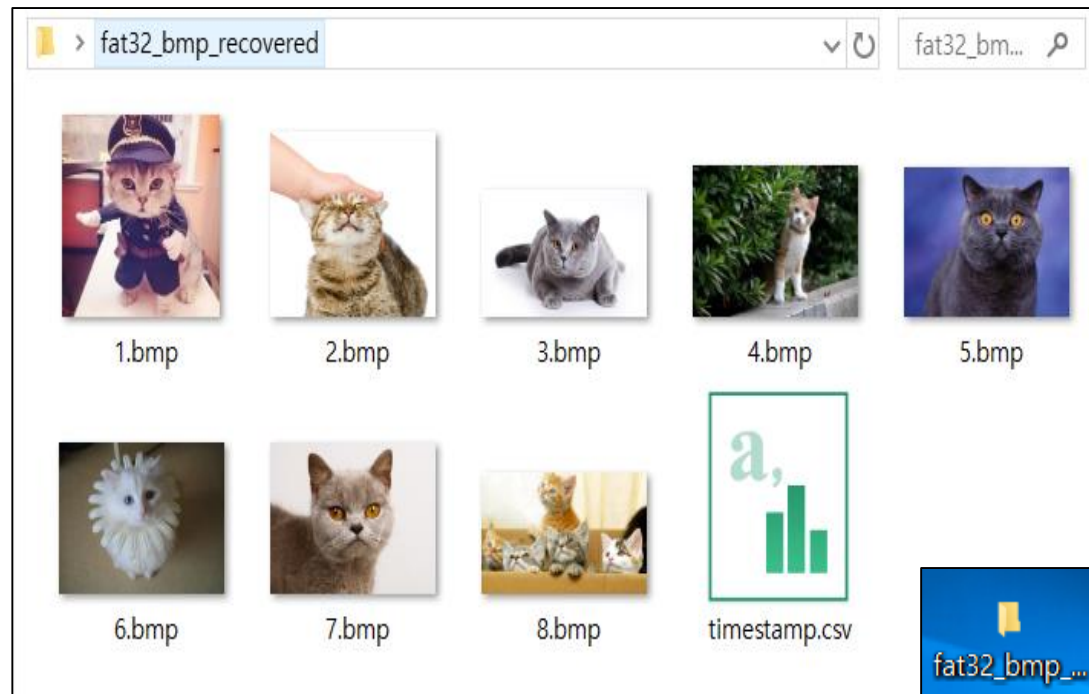
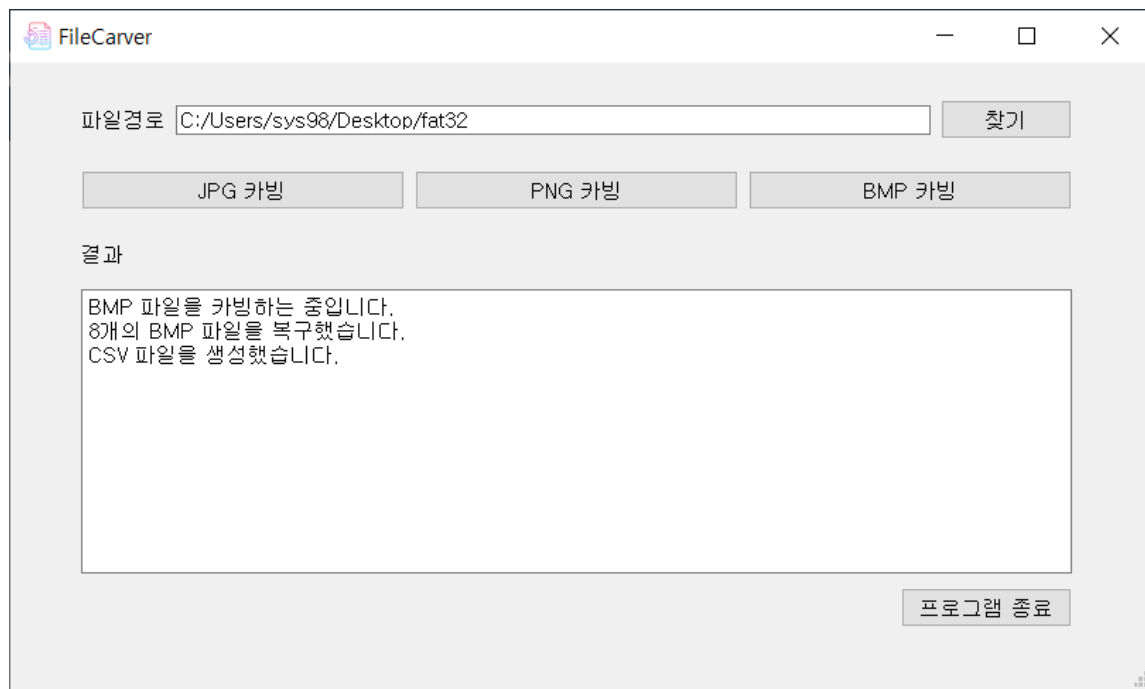
PNG 카빙 알고리즘 | 구현 완료



1. [파일 경로]에서 FAT32 파일시스템을 가지는 덤프 파일[fat32]를 선택한다.
2. [PNG 카빙] 버튼을 클릭하면 PNG 파일을 카빙한다.
3. 결과 화면을 통해 5개의 PNG 파일을 복구했음을 알 수 있다.
4. 1에서 선택한 [파일경로]에 [덤프 파일 이름_png_recovered]라는 폴더가 생기고, 이 폴더에서 이미지를 확인할 수 있다.

02 프로젝트작품 구현 | 제안서에서의 기능과 최종 발표에서의 기능 차이

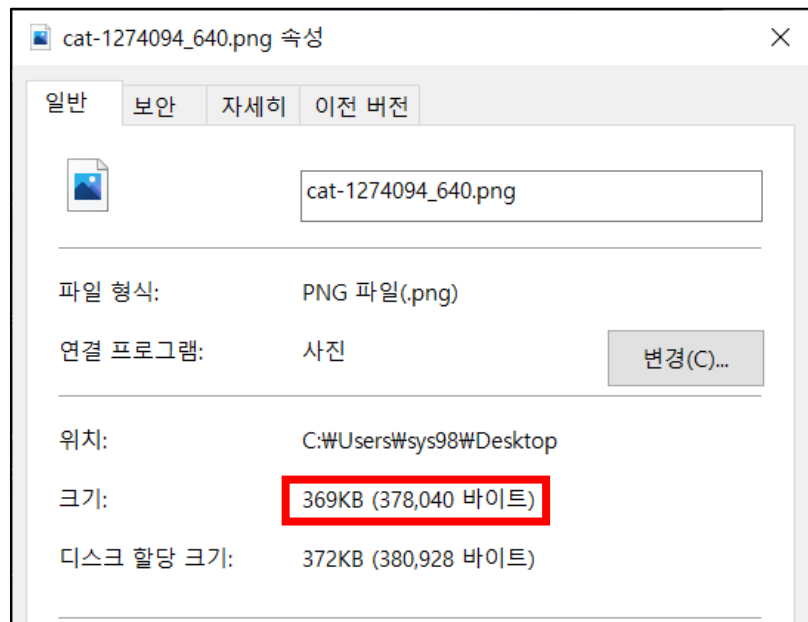
BMP 카빙 알고리즘 | 구현 완료



1. [파일 경로]에서 FAT32 파일시스템을 가지는 덤프 파일[fat32]를 선택한다.
2. [BMP 카빙] 버튼을 클릭하면 BMP 파일을 카빙한다.
3. 결과 화면을 통해 8개의 BMP 파일을 복구했음을 알 수 있다.
4. 1에서 선택한 [파일경로]에 [덤프 파일 이름_bmp_recovered]라는 폴더가 생기고, 이 폴더에서 이미지를 확인할 수 있다.

02 프로젝트작품 구현 | 제안서에서의 기능과 최종 발표에서의 기능 차이

파일 속성정보 추출 | 부분 구현



	A	B	C	D	E	F	G	H
1	Extension	Create Date	Create Time	Last Access Date	Write Date	Write Time	Size	
2	1.png	2017-08-23	22:25:56	2020-05-19	2017-08-19	21:34:46	2618521	b'PNG'
3	2.png	2020-04-29	14:10:16	2020-05-19	2020-04-29	14:10:06	106210	b'PNG'
4	3.png	2020-04-29	14:10:58	2020-05-19	2020-04-29	14:10:56	129377	b'PNG'
5	4.png	2020-04-29	14:11:14	2020-05-19	2020-04-29	14:11:14	378040	b'PNG'
6	5.png	2020-04-29	14:11:48	2020-05-19	2020-04-29	14:11:48	944416	b'PNG'
7								
8								

- 파일 속성정보로는 Create Date, Create Time, Last Access Date, Write Date, Write Time, Size를 추출함
- 1.png의 원본 파일인 cat-1270494_640.png 파일 크기는 378,040byte이고, 코드를 통해 직접 파일 속성정보를 추출해도 378,040byte의 파일 크기가 나온다.

02 프로젝트작품 구현 | 제안서에서의 기능과 최종 발표에서의 기능 차이

파일 속성정보 추출 | 부분 구현

timestamp_t.csv 2020-06-24 오전 7:09 한컴오피스 NEO ...							
	A	B	C	D	E	F	G
1	Extension	Create Date	Create Time	Last Access	Write Date	Write Time	Size
119	117.jpg	2017-08-23	22:24:30	2020-05-19	2017-08-13	20:23:04	315770
120	118.jpg	2017-08-23	22:24:30	2020-05-19	2017-08-13	20:23:04	214924
121	119.jpg	2017-08-23	22:24:30	2020-05-19	2017-08-13	20:25:46	183495
122	120.jpg	2017-08-23	22:24:30	2020-05-19	2017-08-13	20:25:46	173088
123	121.jpg	2017-08-23	22:24:30	2020-05-19	2017-08-13	20:29:18	239960
124	122.jpg	2017-08-23	22:24:30	2020-05-19	2017-08-13	20:29:34	305118
125	123.jpg	2017-08-23	22:24:30	2020-05-19	2017-08-13	20:29:34	163591
126	124.jpg	2017-08-23	22:24:30	2020-05-19	2017-08-13	20:30:36	240156
127	125.jpg	2017-08-23	22:24:32	2020-05-19	2017-08-13	20:30:52	161364
128	126.jpg	2017-08-23	22:24:32	2020-05-19	2017-08-13	20:31:24	270498
129	127.jpg	2017-08-23	22:24:32	2020-05-19	2017-08-13	20:31:24	269261
130	128.jpg	2017-08-23	22:24:32	2020-05-19	2017-08-13	20:32:20	959277
131	129.jpg	2017-08-23	22:24:32	2020-05-19	2017-08-13	20:32:20	1129683
132							
133							

	A	B	C	D	E	F	G
184	183	#####	21:16:18	#####	#####	23:19:34	41565
185	184	#####	21:16:18				
186	185	#####	21:16:18	127	125	125	#####
187	186	#####	21:16:20	128	126	126	#####
188	187	#####	21:19:34	129	127	127	#####
189	188	#####	21:20:51	130	128	128	#####
190	189	#####	21:20:51	131	129	129	#####
191	190	#####	21:20:51	132	130	130	#####
192	191	#####	21:21:00	133	131	131	#####
193				134	132	132	#####
194				135	133	133	#####
195				136	134	134	#####
196				137	135	135	#####
197				138	136	136	#####
198				139	137	137	#####
				140	138	138	#####
				141			

- 왼쪽 사진은 155개의 JPG 파일이 복원되었을 때 CSV 파일에는 129개의 정보만 들어있는 것을 보여줌
- 오른쪽 사진은 덤프 파일 [fat32]에서 모든 JPG의 파일 속성정보를 추출했을 때 191개가 나왔고, 중복을 제외하고 파일 속성정보를 추출했을 때 138개만 얻어졌음을 보여줌

02 프로젝트작품 구현 | 제안서에서의 기능과 최종 발표에서의 기능 차이

파일 속성정보 추출 | 문제점


- 제안서에서는 EXIF 파일에서 파일의 정보(이름, 생성시간, 수정시간, 삭제시간, 파일 크기 등)을 가져오려 함
- EXIF 는 JPG에서만 지원을 하기 때문에 PNG 파일과 BMP 파일에는 적용하지 못함
- Python Image Library인 PIL을 통해 이미지 정보를 불러옴
- BUT 사진을 카빙하는 과정에서 생성/수정/삭제 시간이 카빙한 시간으로 변경되어 본래의 시간정보가 나오지 않음

➡ FAT 32 파일 시스템에서 섹터에 대한 이미지 주소를 인식 후 이에 대하여 시간 정보를 가져옴

- 즉, 작성 시간을 불러올 수 있었음

02 프로젝트작품 구현 | 제안서에서의 기능과 최종 발표에서의 기능 차이

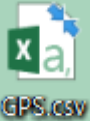
파일 속성정보 추출 | 문제점




KakaoTalk_2020
0624_05111605
1.jpg

KakaoTalk_2020
0624_05124525
1.jpg

KakaoTalk_2020
0624_05142121
6.jpg





GPS.csv



jk.html

NAME	LAT	LON	CTIME	MTIME
KakaoTalk_202	NOT	NOT	2020-06-24 5:11	2020-06-24 5:11
KakaoTalk_202	NOT	NOT	2020-06-24 5:12	2020-06-24 5:12
KakaoTalk_202	37.55	126.9411	2020-06-24 5:14	2020-06-24 5:14

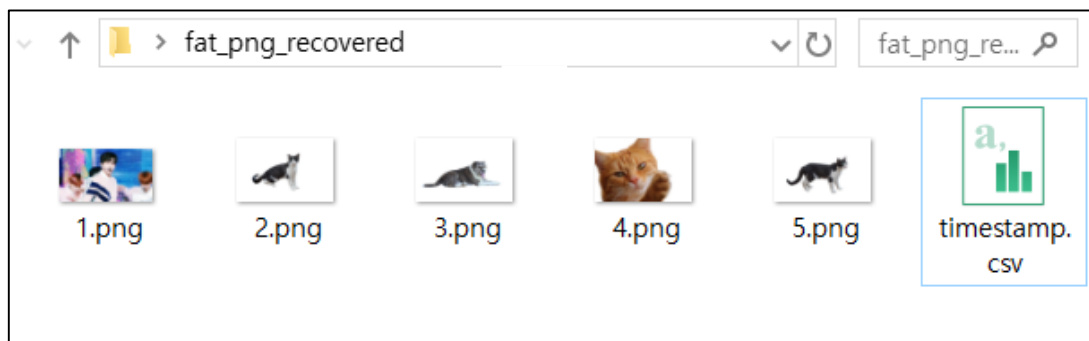


- PIL을 통해서 이미지파일의 정보를 분석 할 땐 위도 경도가 있는 파일에 한하여 csv에 정보를 작성하고 html파일을 생성하여 위치를 표시할 수 있었음
- BUT PIL은 EXIF파일을 분석하는 도구이기 때문에 이를 지원하지 않는 PNG와 BMP에 관한 정보는 나타나지 않음
- 또한 파일 속성정보 추출 섹터분석으로 바꾼 후 시간관계상 jpg파일에 관한 gps정보를 추가하지 못함

02 프로젝트작품 구현 | 제안서에서의 기능과 최종 발표에서의 기능 차이

파일 속성정보 추출 | 문제점

- 파일 이름을 추출하는 과정에서 이름의 길이가 LFN으로 되어있어, 복구 이미지 파일과 실제 파일 이름을 매칭 시키는 것에 어려움을 느낌
- 카빙된 순서대로 번호를 매겨 파일을 식별할 수 있도록 함



	A	B	C	D	E	F	G	H
1	Extension	Create Date	Create Time	Last Access Date	Write Date	Write Time	Size	
2	1.png	2017-08-23	22:25:56	2020-05-19	2017-08-19	21:34:46	2618521	b'PNG'
3	2.png	2020-04-29	14:10:16	2020-05-19	2020-04-29	14:10:06	106210	b'PNG'
4	3.png	2020-04-29	14:10:58	2020-05-19	2020-04-29	14:10:56	129377	b'PNG'
5	4.png	2020-04-29	14:11:14	2020-05-19	2020-04-29	14:11:14	378040	b'PNG'
6	5.png	2020-04-29	14:11:48	2020-05-19	2020-04-29	14:11:48	944416	b'PNG'
7								
8								



02 프로젝트작품 구현 | 제안서에서의 기능과 최종 발표에서의 기능 차이

파일 속성정보 추출 | 배운점

- FAT32 파일의 구조를 알 수 있었다.
- EXIF(메타데이터)에 관한 정보와 EXIF를 지원하는 이미지 파일을 알 수 있었다.
- Python Image library 인 PIL을 통해서 파일 이미지 정보를 추출하는 법을 알 수 있었다.
- Sector에 대한 저장정보에서 필요한 정보를 추출하는 법을 알게 되었다.

02 프로젝트작품 구현 | 제안서에서의 기능과 최종 발표에서의 기능 차이

Out file (CSV파일) | 예상 결과물

	A	B	C	D	E
1	In file name	In file image		Out file name	Out file image
2	a.jpg			A_1.jpg	
3					
4					
5					
6					
7					

*화면 구성이나 옵션 등은 변동가능성 높음

02 프로젝트작품 구현 | 제안서에서의 기능과 최종 발표에서의 기능 차이

Out file (CSV파일) | 구현 완료

파일 이름
생성한 파일을
식별하기 위해 번호 지정

생성날짜 및 시간

접근날짜 및 시간

작성된 날짜

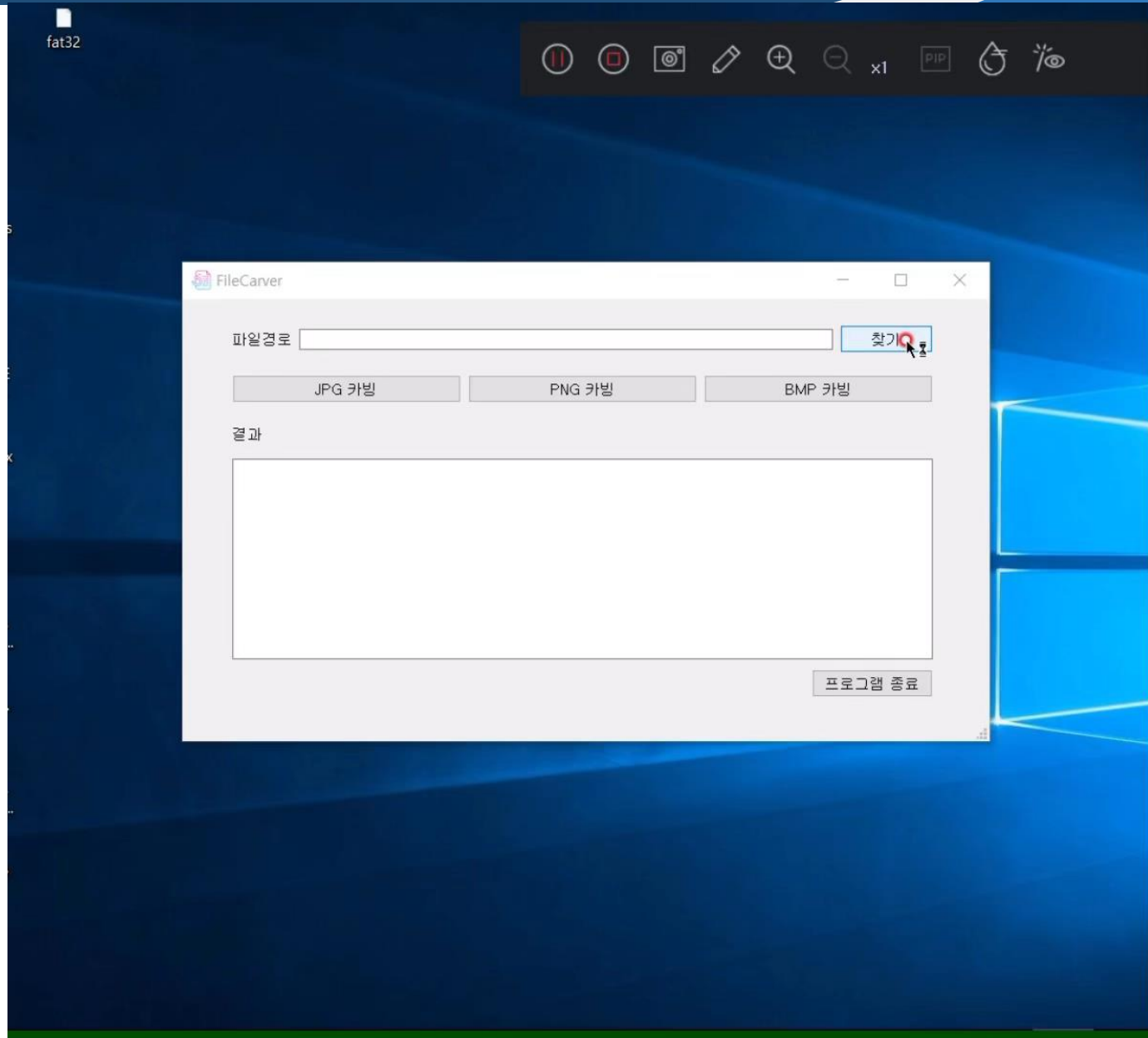
파일 크기

	A	B	C	D	E	F	G
1	Extension	Create Date	Create Time	Last Access Da	Write Date	Write Time	Size
2	1.png	2017-08-23	22:25:56	2020-05-19	2017-08-19	21:34:46	2618521
3	2.png	2020-04-29	14:10:16	2020-05-19	2020-04-29	14:10:06	106210
4	3.png	2020-04-29	14:10:58	2020-05-19	2020-04-29	14:10:56	129377
5	4.png	2020-04-29	14:11:14	2020-05-19	2020-04-29	14:11:14	378040
6	5.png	2020-04-29	14:11:48	2020-05-19	2020-04-29	14:11:48	944416
7							
8							

PNG 파일에 대하여 생성된 CSV파일

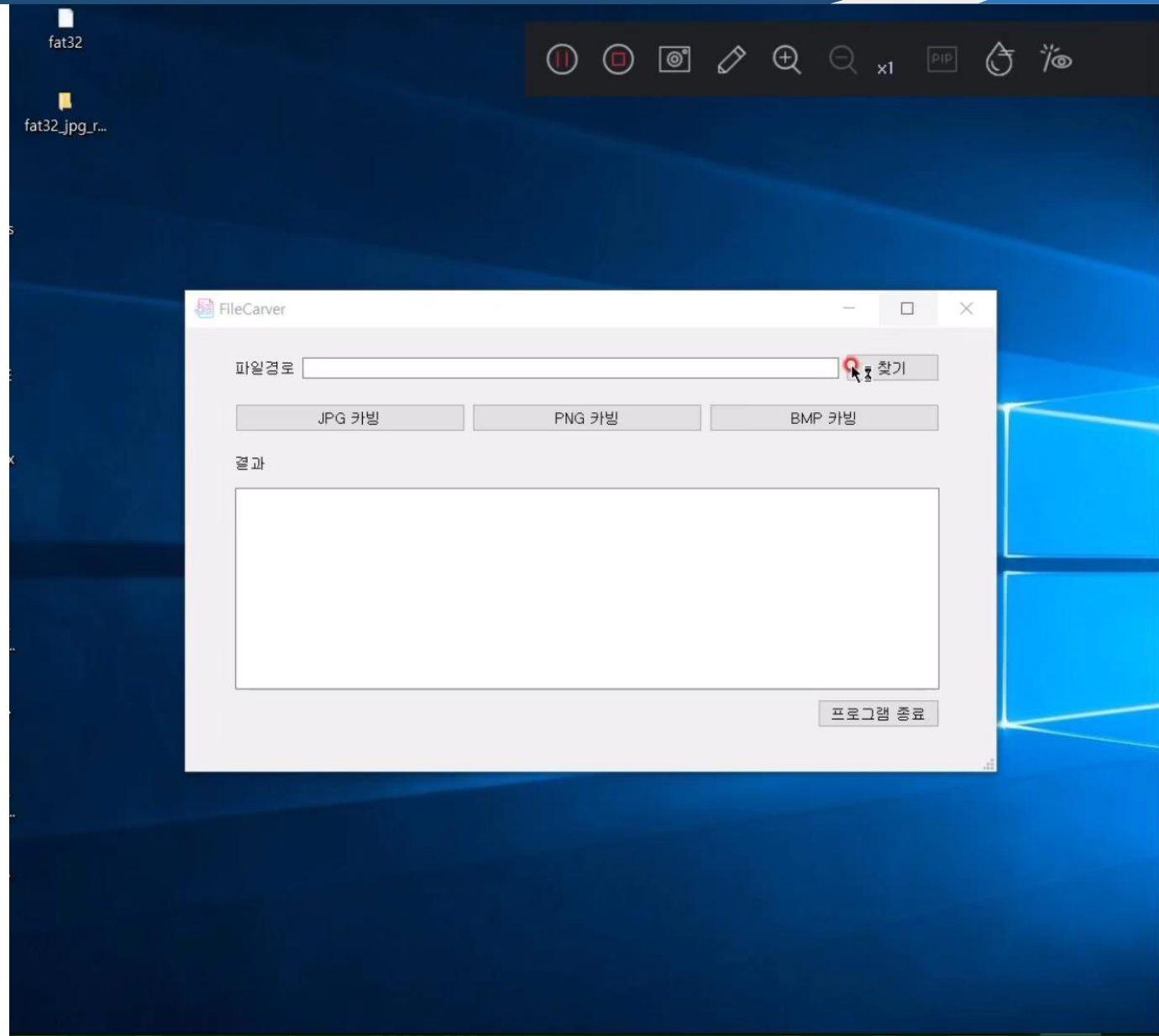
03 프로젝트작품 시연 | JPG file carving

JPG 파일 카빙 시연



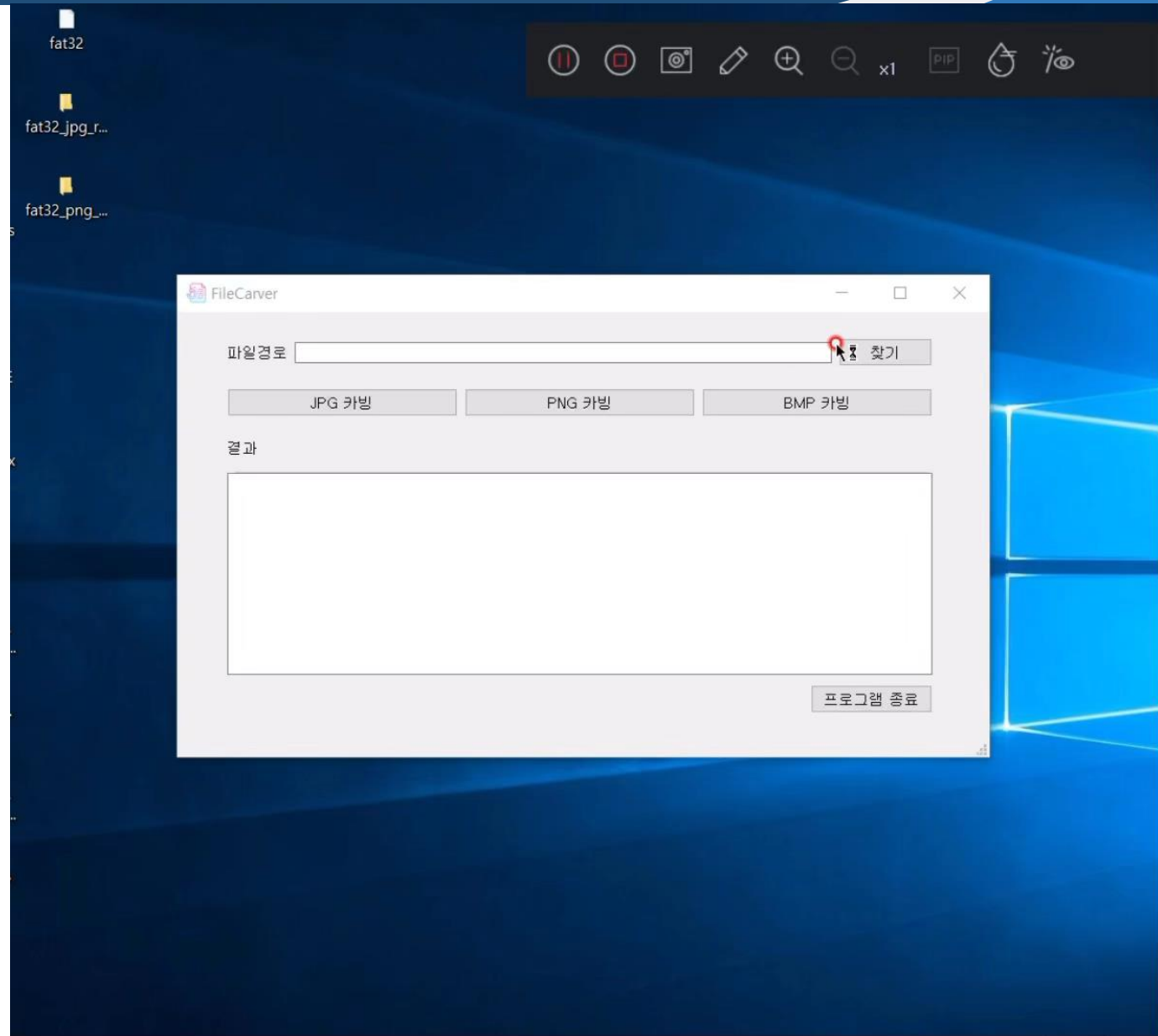
03 프로젝트작품 시연 | PNG file carving

PNG 파일 카빙 시연



03 프로젝트작품 시연 | BMP file carving

BMP 파일 카빙 시연

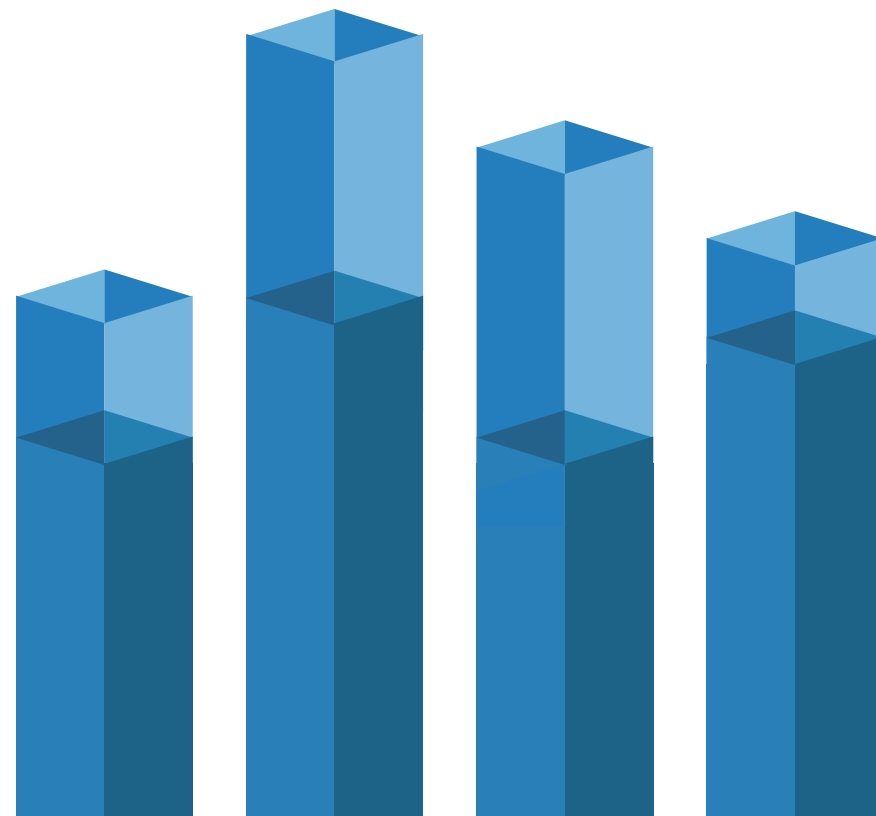


FAT32 뿐만 아니라 다양한 파일시스템 지원

파일속성정보 추출 매칭 + gps정보

CSV 파일에 더 많은 속성 정보 추가

GUI의 개선



- **성연수**
 - FAT32 시스템파일 구조, 파일 카빙 자료조사
 - USB 이미지 덤프
 - 카빙 알고리즘 개발(JPEG, PNG, BMP)
 - Pyqt5로 GUI 개발 및 소스 코드 연결
- **신나연**
 - GUI 구현도구(anaconda, pyQt, PIL) 자료조사
 - 카빙 알고리즘 개발(JPEG, PNG, BMP)
 - 폴더파일 CSV 연결
 - 파일 이름 정보 추출 수정
- **한아름**
 - 이미지 구조 자료(FAT32의 디렉토리 엔트리, JPG, PNG, BMP) 자료조사
 - 카빙 알고리즘 개발(JPEG, PNG, BMP)
 - 이미지 속성정보 추출
 - 시간 변환 프로그램 개발

THANK YOU!