International Journal of Internet, Broadcasting and Communication Vol.13 No.3 124-129 (2021)
http://dx.doi.org/10.7236/IJIBC.2021.13.3.124

IJIBC 21-3-16

Proposal of AI-based Digital Forensic Evidence Collecting System

Eun-Jin Jang[1], Seung-Jung Shin[2]*

[1] *Researcher, Department of IT Convergence, Hansei University, Korea*
dmswls1061@naver.com
[2*] *Professor, Department of ICT Convergence, Hansei University, Korea*
expersin@hansei.ac.kr

Abstract

As the 4th industrial era is in full swing, the public's interest in related technologies such as artificial intelligence, big data, and block chain is increasing. As artificial intelligence technology is used in various industrial fields, the need for research methods incorporating artificial intelligence technology in related fields is also increasing. Evidence collection among digital forensic investigation techniques is a very important procedure in the investigation process that needs to prove a specific person's suspicions. However, there may be cases in which evidence is damaged due to intentional damage to evidence or other physical reasons, and there is a limit to the collection of evidence in this situation. Therefore, this paper we intends to propose an artificial intelligence-based evidence collection system that analyzes numerous image files reported by citizens in real time to visually check the location, user information, and shooting time of the image files. When this system is applied, it is expected that the evidence expected data collected in real time can be actually used as evidence, and it is also expected that the risk area analysis will be possible through big data analysis.

Keywords: *Artificial Intelligence, Digital-Forensic, Evidence gathering, Image File, Location Data*

1. Introduction

With the development of Internet technology, the public has been able to lead a smart life. With the development of Internet technology, the public has been able to lead a smart life. Through this technology, the refrigerator can order milk by itself, and we can control home appliances before coming home. This technology can be realized because data is shared and analyzed for users, and appropriate processing is preceded through analysis. This data continues to be collected, voluminous and smarter through learning. The technology that utilizes the big data learned in this way is called artificial intelligence(AI). Artificial intelligence technology is making rapid progress, and related technologies are already being used in the industrial base. Artificial intelligence technology increases its accuracy through data learning and verification. Therefore, it can be said that it is essential to collect various data.

Therefore, this paper we intends to propose an artificial intelligence-based evidence collection system that analyzes numerous image files reported in real time from citizens to visually check the shooting location,

Manuscript Received: July. 7, 2021 / Revised: July. 12, 2021 / Accepted: July. 15, 2021
Corresponding Author: expersin@hansei.ac.kr
Tel: +82-31-450-9585, Fax: +82-31-450-5172
Professor, Department of ICT Convergence, Hansei University, Korea

user information, and shooting time of the image files. If this system is used, it is expected that it will become a new evidence collection device that can visually check the location information in a situation where there is a suspicion of a crime by analyzing the location information collected. In addition, it is expected that correlation analysis of dangerous areas will be possible through the analysis of location data collected using artificial intelligence technology.

2. AI Technology

2.1 The techniques of Data Analysis

Data is a very important component in artificial intelligence technology. This is due to the increased utilization of the data by finding data associations and representing them visually. The data obtained using several methods are analyzed through the following analytical techniques. Data analysis techniques include visualization, spatial analysis, exploratory data analysis, statistical analysis, and data mining [1]. Visualization is the most basic analysis method and is usefully used in exploratory analysis methods. Spatial analysis visually represents space and related data, and exploratory data analysis derives meaningful data through values of various dimensions. Statistical analysis refers to numerical and pictorial analysis to comprehensively understand a specific phenomenon, and data mining extracts valuable information by analyzing data of specific patterns and rules from large-scale data [4]. Table 1 shows the types and characteristics of data analysis techniques.

Table 1. Types and Characteristics of Data Analysis Techniques

data analysis technique	Characteristic
Visualization	Useful in exploratory analysis
Spatial analysis	Visualizing spatial data and spatial related data
Exploratory data analysis	Deriving meaningful data through values of various dimensions
Statistical analysis	Numerical and pictorial analysis to comprehensively identify specific phenomena
Data mining	Extract valuable information by analyzing data of specific patterns and rules from large-scale data

3. Digital Forensic Evidence Gathering Process

Digital forensic investigation procedures include preliminary preparation, evidence collection, packaging and transport, investigation analysis, detailed review, and report writing. Figure 1 shows the digital forensic investigation procedure.

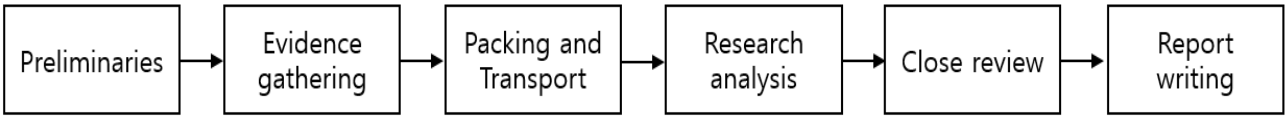


Figure 1. Digital Forensic Investigation Process

Prompt collection of evidence is an essential factor for an accurate investigation. However, in the current situation, procedures such as tool preparation, tool verification, and analysis training are required in the pre-preparation process before the evidence collection process. Also, a warrant must be issued to collect evidence. In a forensic investigation situation in which evidence must be collected in a timely manner, there may be a risk of evidence destruction if the existing evidence collection procedure that requires a series of procedures is used.

If you look at the criminal cases disclosed to the media, even when there are suspicions, the charges are often not proven due to insufficient evidence, and there are cases where the victims are unfairly caused by the unclear evidence. However, if a specific person intentionally damages the evidence or the evidence is lost after a long period of time, it is difficult to secure the whole evidence. Therefore, it can be said that there is a need for a new evidence collection system that can be used as clear evidence in the investigation process

Therefore, in this paper, we propose an AI-based evidence collection system that can visually check the shooting location, user information, and shooting time of image files by analyzing numerous image files reported by citizens in real time.

4. Proposed AI-based Digital Forensic Evidence Collecting System

The evidence collection system proposed in this paper is a system that analyzes image files received through citizens' reports, stores user information, location information, and shooting time in the evidence collection DB, and visually displays the analyzed location information on the map. By converting the collected data into a DB, the investigator analyzes the location information stored in the photos to visualize the data, and converts the related information into a DB and stores it in the evidence management DB. The data stored in this way will notify users through correlation analysis of risk areas in the future, so that the effect of crime prevention can also be expected. Figure 2 shows the flow chart of the evidence collection system proposed in this paper.

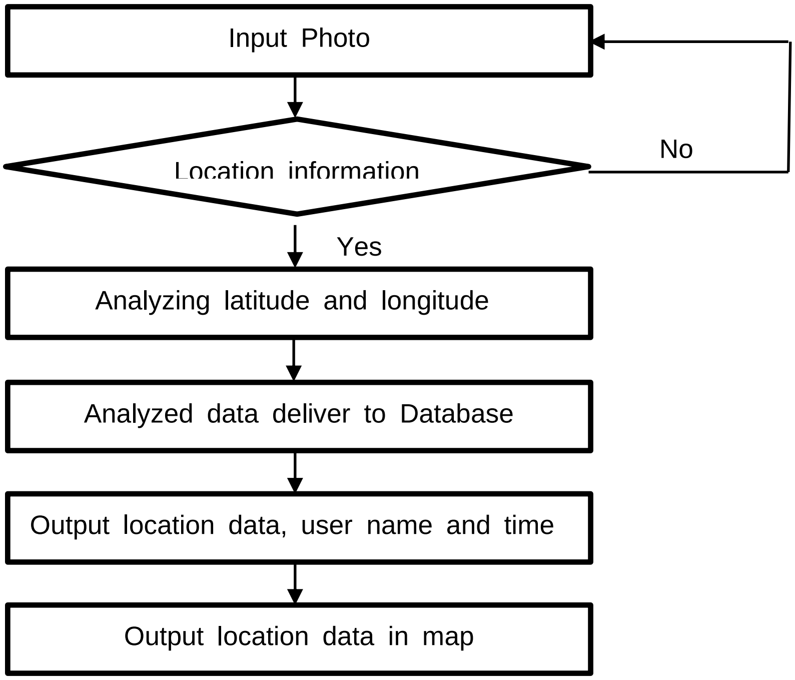
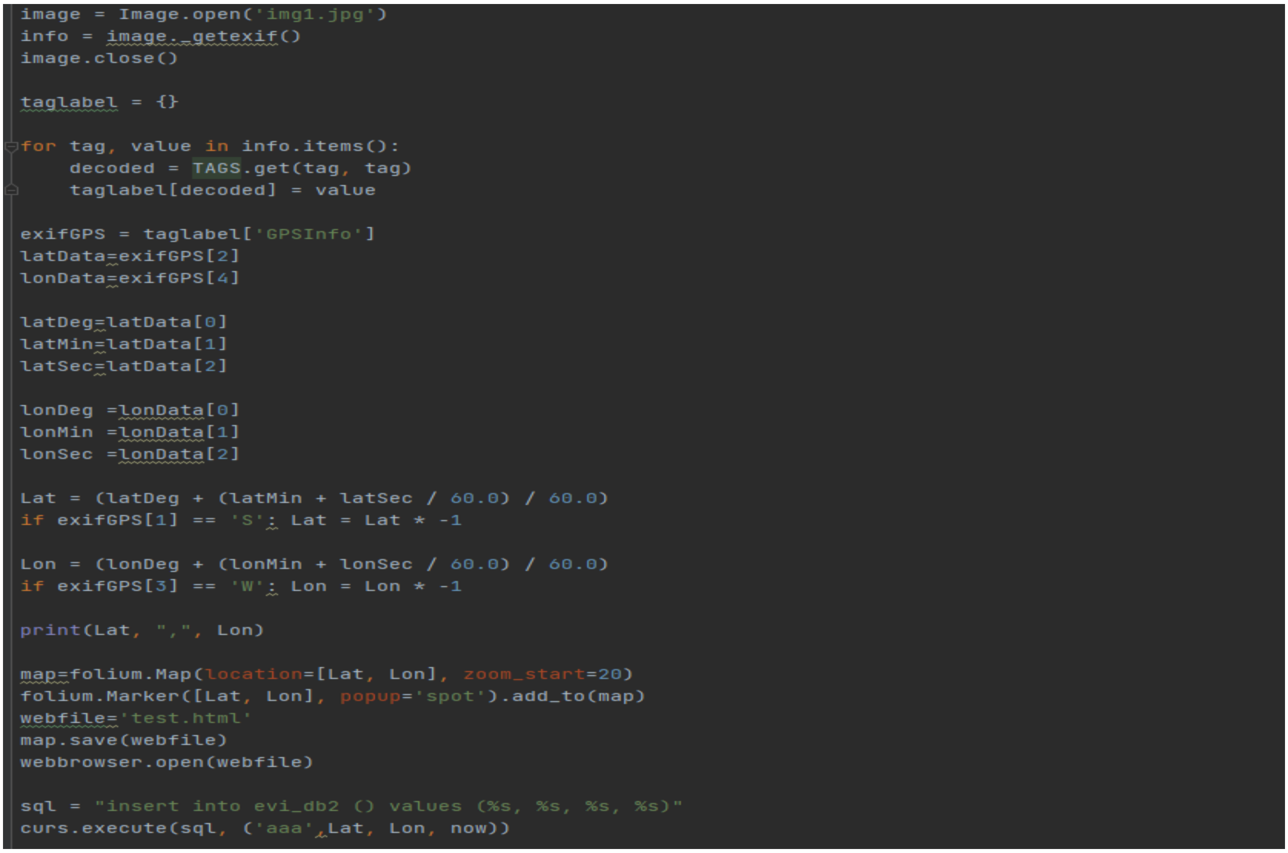

Input Photo
Location information
No
Yes
Analyzing latitude and longitude
Analyzed data deliver to Database
Output location data, user name and time
Output location data in map

Figure 2. Flowchart of the Evidence Collecting System

The development environment for implementing the system proposed in this paper is as follows. The OS environment is Windows 64bit, and the development language is implemented through Python. The development tool used Pycharm, and the database tool used MySQL. Figure 3 is the reported photo. Figure 4 shows the code of the evidence collection system.

Figure 3. Reported Photo



```
image = Image.open('img1.jpg')
info = image._getexif()
image.close()

taglabel = {}

for tag, value in info.items():
    decoded = TAGS.get(tag, tag)
    taglabel[decoded] = value

exifGPS = taglabel['GPSInfo']
latData=exifGPS[2]
lonData=exifGPS[4]

latDeg=latData[0]
latMin=latData[1]
latSec=latData[2]

lonDeg =lonData[0]
lonMin =lonData[1]
lonSec =lonData[2]

Lat = (latDeg + (latMin + latSec / 60.0) / 60.0)
if exifGPS[1] == 'S': Lat = Lat * -1

Lon = (lonDeg + (lonMin + lonSec / 60.0) / 60.0)
if exifGPS[3] == 'W': Lon = Lon * -1

print(Lat, ",", Lon)

map=folium.Map(location=[Lat, Lon], zoom_start=20)
folium.Marker([Lat, Lon], popup='spot').add_to(map)
webfile='test.html'
map.save(webfile)
webbrowser.open(webfile)

sql = "insert into evi_db2 () values (%s, %s, %s, %s)"
curs.execute(sql, ('aaa',Lat, Lon, now))
```

Figure 4. Code of the Evidence Collecting System

Figure 5 shows the photography location information identified through this system. Also, Figure 6 shows the location and user information delivered to the evidence collection DB.

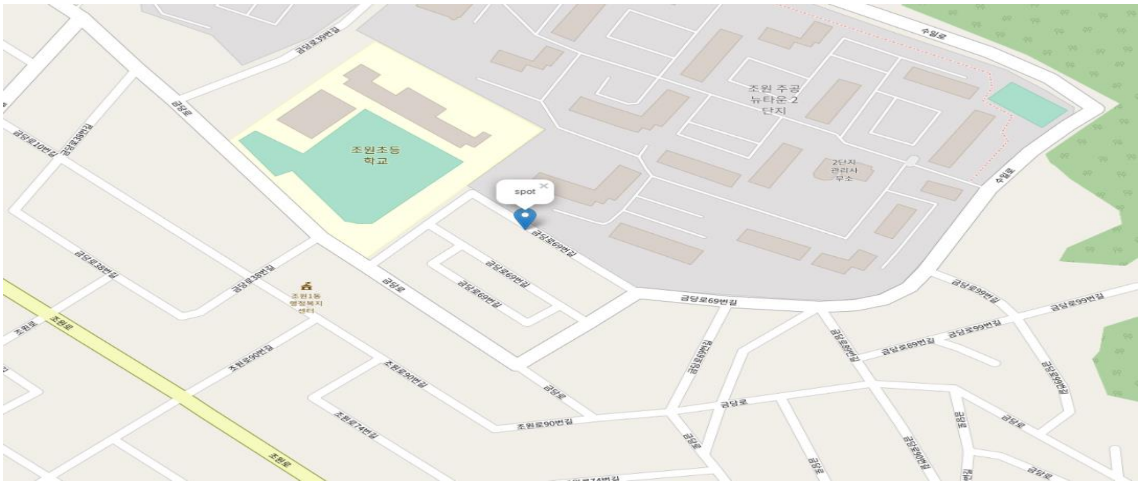
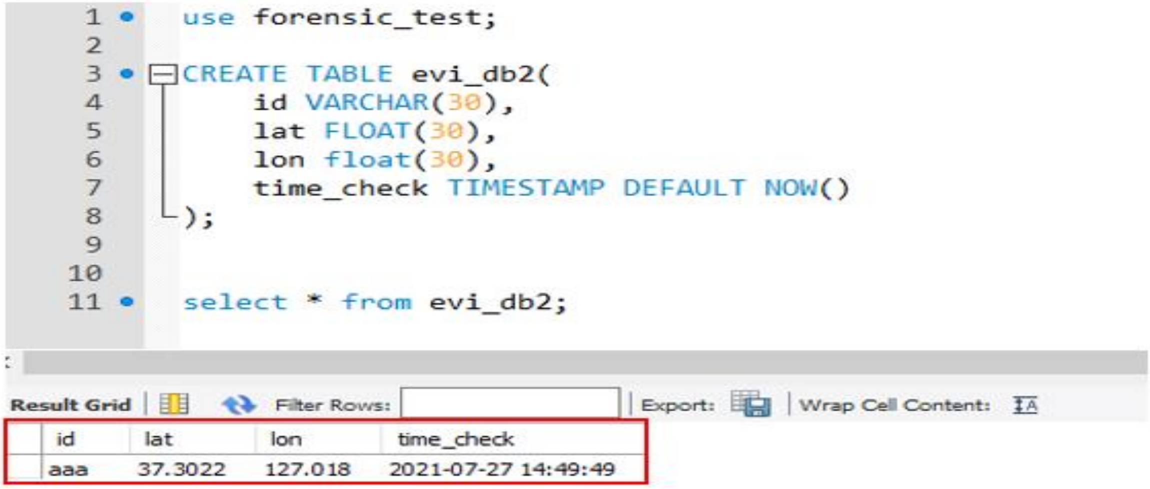


Figure 5. Photo Location Information verified through the System



```
use forensic_test;

CREATE TABLE evi_db2(
    id VARCHAR(30),
    lat FLOAT(30),
    lon float(30),
    time_check TIMESTAMP DEFAULT NOW()
);


select * from evi_db2;
```

id	lat	lon	time_check
aaa	37.3022	127.018	2021-07-27 14:49:49

Figure 6. Location and User Information passed to the Evidence Collecting DB

5. Conclusions

In this paper, we proposed an AI-based evidence collection system that can efficiently collect evidence in the digital forensic investigation process. It aims to collect evidence that can prove the suspicion of a crime from the image files reported in real time. If this system is used, it is expected that it will become a new evidence collection device that can visually check the location information in a situation where there is a suspicion of a crime by analyzing the location information collected. In addition, it is possible to analyze the correlation of the collected location data by transferring the photo information to the linked DB, and it is expected that it will be able to be used as a predictive data for crime prevention in dangerous areas. Since the system proposed in this paper targets the JPG file format, analysis techniques for various file formats will need to be added to collect more diverse evidence. In addition, due to the nature of this system that handles user information, additional research on security should be conducted.

References

[1] Jiyoon Ham, Joshua I. James, “A Feature Comparison of Modern Digital Forensic Imaging Software”, The Journal of The Institute of Internet, Broadcasting and Communication (IIBC), Vol. 19, No. 6, pp. 15-20, Dec 2019
DOI: https://doi.org/10.7236/JIIBC.2019.19.6.15

[2] Eunjin Jang, Seungjung Shin, “Proposal of New Data Processing Function to Improve the Security of Self-driving Cars System”, The Journal of The Institute of Internet, Broadcasting and Communication, Vol.20, No.4, pp.81-86, Aug 2020
DOI: http://doi.org/10.7236/JIIBC.2020.20.4.81

[3] Changseob Yun, Jaehyun Jun, Sungho Kim, Daesoo Kim, “Real-Time Forensic Marking Method Based on Multi-Core”, The Journal of Korean Institute of Communications and Information Sciences, Vol. 45, No. 01, pp. 212-221, Jan 2020
DOI: https://doi.org/10.7840/KICS.2020.45.1.212

[4] Jinseong Park, Seunghee Seo, Yeog Kim, Changhoon Lee, “A Study of the Decrption Method of LockMyPix’s Media Files for Forensic Analysis”, The Journal of Digital Forensics, Vol. 14, No. 3, pp. 269-278, Sep 2020
DOI: https://doi.org/10.22798/KDFS.2020.14.3.269

[5] Jaejeong Hwang, “Forensic Detection of Filteration Forgery in Digital Images”, The Journal of The Institute of Electronics and Information Engineers, Vol. 56, No. 1, pp. 85-91, Jan 2019
DOI: https://doi.org/10.5573/IEIE.2019.56.1.85

[6] Jeewon Jang, Seunggyu Bang, Jaehyeok Han, Sangjin Lee, “A Recovery Technique of PDF File in the Unit of Page”, The Journal of The Korea Information Processing Society, Vol. 6, No. 1, pp. 25-30, Jan 2017
DOI: https://doi.org/10.3745/KTCCS.2017.6.1.25

[7] Haejin Lee, Taeshik Shon, “E-mail Header-Based Search and Seizure for Internet Portal Digital Forensics”, The Journal of The Korea Institute of Information Security & Cryptology, Vol. 28, No. 5, pp. 1129-1140, Oct 2018
DOI: https://doi.org/10.13089/JKIISC.2018.28.5.1129

[8] Sanghuk Yoon, Sangjin Lee, “A Study on Digital Evidence Automatic Screening System”, The Journal of Digital Forensics(KDFS), Vol.14, No.3, pp.239-251, Sep 2020
DOI: http://doi.org/10.22798/KDFS.2020.14.3.239

[9] Sojung Oh, Taegi Lee, Gibum Kim, “A Study on Digital Forensic Image Production Model”, The Journal of Digital Forensics(KDFS), Vol.15, No.2, pp.137-150, Jun 2021
DOI: http://doi.org/10.22798/KDFS.2021.15.2.137

[10] Yangsub Kwon, “A Study on Korean Digital Forensic Investigation Procedure and Construction of Verification System”, The Journal of Digital Forensics(KDFS), Vol.15, No.1, pp.67-82, Mar 2021
DOI: http://doi.org/10.22798/KDFS.2021.15.1.68