https://doi.org/10.2352/EI.2025.37.3.MOBMU-314

Automated Monitoring of Stolen Cultural Artifacts on Online Marketplaces

Huajian Liu, York Yannikos, Julian Heeger, Simon Bugert, Waldemar Berchtold, Martin Steinebach
Fraunhofer Institute for Secure Information Technology (SIT), Darmstadt, Germany

Abstract

Tracking and identifying stolen cultural artifacts on online marketplaces is a daunting task that has to be accomplished through manual search. In this paper, an automated monitoring tool is developed to track and identify stolen cultural goods on targeted online sales platforms. In case of theft, the original owner can upload descriptive keywords and photos of the stolen objects to start monitoring tasks to track and identify the stolen objects on targeted online marketplaces and get alerted when identical or highly similar objects appear on the monitored sales platforms. The technical challenges posed by automated monitoring are addressed by proposed advanced crawling and image feature extraction and matching solutions. With the support of proposed novel techniques, the developed monitoring tool can efficiently and effectively monitor stolen artifacts on online marketplaces, significantly reducing the manual inspection effort.

Introduction

Illicit trade in cultural artifacts has become a source of financing for criminal organizations [1]. Many of them are stolen artifacts. Museums and collectors frequently have their cultural assets stolen, which sometimes results in a loss of millions. Thieves usually want to exchange the stolen goods for money. To this end, stolen artifacts are offered for sale through various channels, including online sales platforms, to attract as many potential buyers as possible [2][3].

Therefore, monitoring these sales platforms becomes an important means of tracking the whereabouts of stolen artifacts. However, manual search for stolen artifacts on various websites is not only time-consuming but also cost-intensive, because it must be conducted by trained personnel. Moreover, given the vast number of items and rapid updates on online sales platforms, manual searches can easily miss detecting some items or their updates.

The objective of this work is to develop a software solution that enables owners of cultural goods as well as law enforcement agencies to track and identify cultural goods on online sales platforms in case of theft. The entire monitoring process should be highly automated, which can search for stolen objects based on descriptive keywords and object photos. Different targeted online marketplaces should be scanned at the same time and the information on relevant offerings should be automatically collected and stored.

However, keyword-based searches will inevitably result in many items with close properties appearing in the search results, especially when the objects to be searched for belong to a common category. To exclude irrelevant items from the search results and minimize the workload of subsequent manual inspection, the crawled items should be visually compared to the stolen objects and only items with identical or highly similar appearance will be finally listed in the results.

To achieve the above-mentioned objective, in this work we developed a system named KIKu-Mon for automated monitoring of stolen objects on online marketplaces. Through a user-friendly web interface, users can easily manage stolen artifacts and set up customized monitoring tasks. Once suspect items are found, the user will be alerted, and the found items will be presented to the user for inspection. The proposed monitoring system is oriented towards practical applications, which is designed and developed for real-world scenarios.

To automatically collect relevant offerings from various websites, we developed novel advanced crawling tools to fetch data, which include web crawlers and data storage schemes. The developed crawler is driven by recipes described in domain specific language (DSL) and is quickly adaptable and patchable via updating recipes in case of change of website layout.

To visually compare the appearance of crawled items with stolen artifacts and efficiently remove irrelevant items, we developed image matching mechanisms specifically tailored for recognition of cultural artifacts. Deep learning models are fine-tuned using specific large artifact datasets to extract image features.

This paper is organized as follows. Section 2 introduces the proposed KIKu-Mon system and its components. Section 3 discusses the technical challenges and proposes the corresponding solutions. Section 4 presents the evaluation results. The paper is concluded in Section 5.

Proposed KIKu-Mon System

As shown in Figure 1, the proposed KIKu-Mon system is designed as a client-server system, which consists of a web application as the client and a set of advanced crawling and image matching tools as the server.

The web application serves as the user interface to the system on the client side, which runs on smartphones, tablets and PCs, as shown in Figure 2. In the webapp, users can upload and manage stolen artifacts and submit monitoring tasks. To each stolen artifact, users can add photos, metadata, and descriptive keywords. The keywords describe the artifact attributes and will be used in the search for suspect artifacts on every target marketplace. The added photos are used for visual comparison with the crawled items to identify those that are highly like the stolen artifact.

When setting up a monitoring task, users can specify the stolen artifacts to be monitored, select the marketplaces to be monitored, and set the monitoring frequency and duration. In each task, multiple stolen artifacts and multiple marketplaces can be selected. A task can be started immediately or at a specified time. If a task is set to be recurrent, it will be automatically and repeatedly started at the specified interval, e.g. every two days, every week or every month, till the end date. All tasks are submitted to the server and performed on the server.

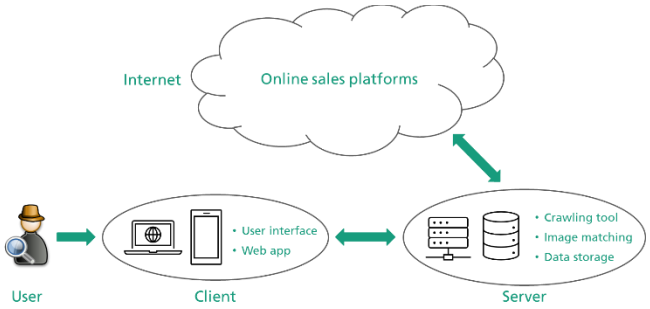


***Figure 1**. KIKu-Mon system*

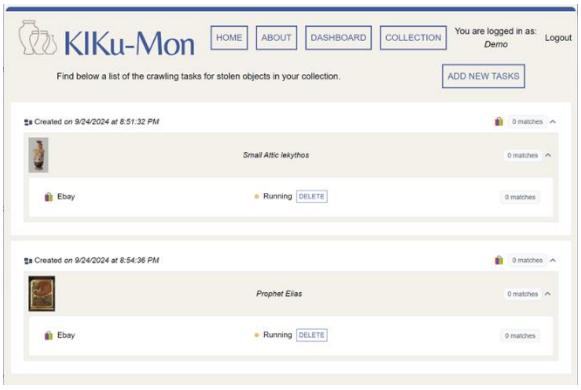


***Figure 2.** KIKu-Mon web application*

The monitoring results, generated on the server, will be transferred to the client and presented to users through the webapp. The results are categorized by stolen artifacts and marketplaces, presenting which items are found on which marketplaces. For convenient inspection, it is also shown how a found item is matched with the stolen artifact, for instance, which photos of the item is matched with the stolen artifact with the highest similarity score.

The core tasks of KIKu-Mon are performed on the server-side. The server not only hosts the core functions of KIKu-Mon, including the advanced crawling tools and the image matching tool, but also handles the data storage. Upon receiving a submitted monitoring task from the client, the crawling tools will start searching for the submitted stolen object on the selected online sales platforms based on the given keywords. The information of the found items including their images will be downloaded and temporarily stored in a data pool on the server. Subsequently, the crawled objects will be visually compared with the submitted stolen artifacts by the image matching tool. Only items with identical or highly similar appearances will be transmitted back to the client and shown to the user as results, which are sorted by their similarity scores.

KIKu-Mon is an asymmetric system that requires significantly higher performance from the server than from the client. Not only does the server need to handle tasks with higher computational demands, but the server must also be able to process multiple monitoring tasks submitted by different clients in parallel. Depending on the volume of monitoring tasks, the server will need to be scaled up accordingly to provide adequate performance. The server does not necessarily have to be a single physical computer but can also be a cluster of intelligently managed computers if required.

Technical Challenges and Solutions

Automated monitoring of stolen artifacts on different online marketplaces poses two main challenges: collection of relevant offerings from online sales platforms and identification of stolen objects. For each of them, we have developed new targeted technical solutions, respectively.

Advanced Crawling Tools

For collection of relevant items for sale, we developed advanced crawling tools to gather items from websites of different online sales platforms [4]. We focus on using the search engine of the websites of marketplaces which typically gives sufficiently precise results. Many websites allow their search results to be filtered according to certain properties, which reduces the number of resulting items to be crawled.

Our advanced crawling architecture consists of three components as shown in Figure 3: data storage, REST API and crawlers.

1. **Data Storage:** The data storage component consists of a database and a file system storage where the results of the individual crawlers are stored. The database can be accessed via a REST API component to enforce proper authorization and logging for the individual crawls. While text data is stored directly in the database, larger media files such as images or other binary data are stored in the file system storage so that they can be accessed and deduplicated efficiently.
2. **REST API:** The REST API component manages the communication between the other components and provides an interface to fetch new crawl tasks from an integrated task queue, submit crawl results (i.e. store crawled data), and read back stored data if necessary. Crawlers typically submit new data to the database via the REST API, while the API taking care of authentication, validation and deduplication. The new data from a crawler can result in new tasks (e.g. newly discovered subpages of a website with additional items to be crawled) which are submitted to the REST API and put on the task queue to be distributed to the crawlers.
3. **Crawlers:** The crawlers are independent components that can run on different devices or virtual machines and are built to have just enough knowledge to understand and use a domain specific language (DSL) we developed for our crawling architecture. They maintain a connection with the REST API to communicate back their status (e.g. in order to communicate or receive information about possible network problems).

In order to divide complex crawling tasks into smaller, reusable subtasks, we have developed a DSL. Each small subtask, such as downloading an image based on a URL or hashing a binary blob with a specific algorithm, is described in our DSL and builds a so-called recipe that can be reused in other larger crawling tasks. This allows us to represent typical larger and more complex crawler tasks as a collection of recipes and enables us to quickly adapt and change crawling tasks if issues arise, such as a changed website layout. Recipes can be updated by a developer and are versioned and stored in the database. When crawlers take on tasks, they collect the latest changes to all included recipes from the

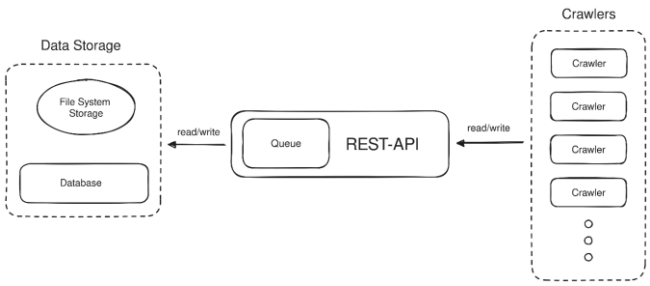


Figure 3. *Advanced crawling architecture*

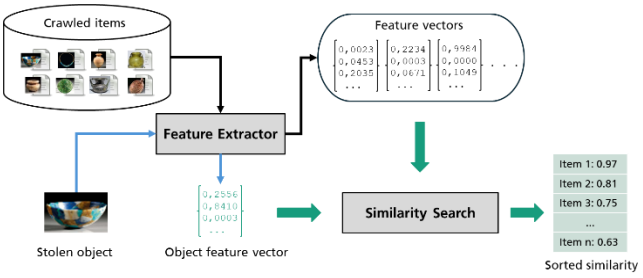


Figure 4. *Visual matching of the stolen artifact and the crawled items*

database via the REST API. This allows us to patch and modify crawlers while they are active without the need to modify their source code.

Image Matching Tool

Since a keyword-based search is used when searching sales platforms, all items that match keyword descriptions will be found and downloaded. The range of search results therefore depends greatly on the keywords used.

Because stolen artifacts are likely to be sold with non-specialized text describing the artifacts which differs from the official description, the keyword descriptions used for searching should not be too precise as this would greatly increase the probability of a missed hit. Using fuzzy keyword descriptions will maximize the chance that the target items will appear in the search results, but this will also lead to a significant increase in the number of search results, allowing many items with close properties to be included in the search results. If used keywords are even inaccurate, it can lead to a large number of irrelevant items appearing in the results. In short, a keyword-based search alone can not prevent many close or even irrelevant items from appearing in search results, especially when the objects to be searched for belong to a common category.

In order to exclude irrelevant items from the search results and to identify items with the same or highly similar appearance to the stolen artifacts to be monitored, deep learning-based feature extraction and image matching techniques have been developed to visually match the appearance of the crawled items with the stolen artifact.

As shown in Figure 4, our image matching tool consists of two main components: a feature extractor and a similarity search module. The feature extractor is used to extract feature vectors from the photographs of the stolen object and the crawled items, which can be a classical handcrafted feature extractor or a neural network model pre-trained for feature extraction. Since the photographs of a stolen object posted on an online marketplace could be different from the ones officially documented in the museum database, it is required that the extracted features from different photographs of the same object should share high similarity, even if the photographs vary in terms of resolution, visual quality, shooting angle, orientation, background, etc.

First, a feature vector is extracted from the photograph of the stolen object to be monitored. In the same way, a pool of feature vectors is created by extracting a feature vector from each photograph of a crawled item. Subsequently, the similarity search module evaluates the visual similarity between each crawled item and the stolen item by comparing the similarity of their feature vectors.

Items with low similarity will be filtered out, and only items with high similarity will be retained and shown to the user as matched objects. In this way, the probability of false alarms can be significantly reduced, thus greatly reducing the workload of subsequent manual inspection procedures.

Unlike general object recognition, the identification of visually similar archaeological objects poses new challenges. First, archaeological objects possess unique characteristics, which depend on their period and geographical origin. Second, the crawled items all match the given keyword descriptions used for searching, so they belong to the same category in one way or another, which often share similar appearances like shape, color, texture, pattern, decoration and so on.

This requires that the image matching tool should be able to distinguish between archaeological artifacts belonging to the same category but possibly originating from different eras and regions. However, general deep learning models, e.g. ResNet, EfficientNet, etc., are trained for common objects and not optimized for cultural goods.

In order to extract appropriate features required for the identification of archaeological objects, we apply the CNN model MaxAvgCat proposed in [5] as the feature extractor, which is developed and trained specifically for recognition of cultural artifacts. MaxAvgCat uses ResNet as backbone and is fine-tuned through transfer learning [6] with large datasets of archaeological objects. Moreover, it extends ResNet with a new head so that it does not only output high-level features but can extract multi-level features, which better represent the unique characteristics of archaeological objects. Thus, image matching based on such features can distinguish similar objects of the same category more accurately.

Privacy Protection

The data storage implements privacy protection by applying the following security and data protection measurements:

1. All storage media (e.g. HDDs, SSDs) used for data storage are encrypted and suitable access roles are enforced such that no unauthorized third party can access the data.
2. All crawled data (e.g. items with images) are stored only for a short period of time, i.e. until the newly crawled images are checked against the photos of the stolen objects using the image matching tool. This is done in regular intervals (e.g. daily at midnight) using a scheduled job that compares photos of the stolen artifacts from users to those of the corresponding crawled items. If no match is found for a crawled item, its data is deleted from the database and file storage. If a match is found, the relevant data about the matched item is returned to

the user for review and validation. Accordingly, any crawled data is stored for a maximum of one day.

Evaluation Results

In the current version of KIKu-Mon, crawlers for six websites of different kinds of online marketplaces have been implemented, including Catawiki [7], eBay [8], etc.

Test Setup

KIKu-Mon has been tested using different kinds of artifacts from two museums in a near-real operating environment. In a simulated theft scenario, stolen artifacts are re-photographed to obtain photographs that differ from the documented official ones. The re-photographed photographs differ from those in the official record in various aspects, such as shooting angle, lighting and resolution. The stolen objects are listed for sale on online marketplaces, using either high- or low-resolution photos and text descriptions that differ from the official descriptions.

Test Results

The owner starts monitoring tasks in KIKu-Mon to search for the stolen objects on selected marketplaces, using the official professional high-resolution photographs and the official text descriptions provided by museums. A few keywords for each object are extracted from the artifact's official description and metadata. All items matched with the chosen keywords are downloaded by the crawlers and subsequently visually compared with the uploaded official photographs of stolen objects using the image matching tool. The number of matched items for each stolen artifact varies greatly, depending on the number of similar items available on the market. Finally, the items with similarity scores higher than a predefined threshold are presented as results.

The image matching tool based on the features extracted by the dedicated CNN model for artifact recognition achieved sufficient accuracy to exclude irrelevant items and identify the same and highly similar objects. In all cases, items with dissimilar appearance were successfully filtered out and the correct items were successfully identified and listed in the results. In terms of the accuracy of similarity scores, the features extracted by the specialized model for cultural artifacts outperformed those extracted by the original ResNet.

It is noticed that in all tests with sharp item photographs, the correct stolen objects are correctly identified with high similarity score. The blurriness of the item photographs is shown to have a significant impact on the matching results, while the resolution of download images has little effect on the matching results.

Crawler Performance

To evaluate the crawler performance, we collected data of about 9,200 items in four different categories from Catawiki in weekly crawling runs. On average, the crawlers required about 1.6 seconds to collect one item, where most of the time was spent on creating a screenshot of the item webpage. The rest of the crawling runs were significantly faster, e.g. about 50 image files can be downloaded per second.

In another test, the Coins category on Catawiki was crawled and only the HTML content was saved. Out of a total of 858 search entries, it took 176 seconds, on average 205 milliseconds per search entry. When creating screenshots and saving all associated images, 50 search entries took 75 seconds, i.e. 1.5 seconds per entry. This result is consistent with the results from weekly runs.

Conclusion

In this work we proposed an automated monitoring tool, KIKu-Mon, for tracking and identifying stolen cultural goods on online marketplaces, which runs in a browser on PCs or mobile devices and does not require any pre-installed software. With KIKu-Mon users can start customized tasks to monitor selected online sales platforms, which can greatly reduce the manual search and inspection effort. Two challenges posed by automated monitoring, including collecting relevant offerings and identifying stolen objects, are addressed by novel technical solutions. An advanced crawling architecture is developed for data collection from online sales platforms, in which a domain specific language (DSL) is developed to enable efficient crawling and allow easy adaptation of crawlers for website changes. A dedicated CNN model specialized for artifact recognition is used to extract multi-level features to identify identical or highly similar artifacts.

In future work, more functions will be developed in the next version. For example, users can customize the scope of the keyword-based search by selecting either fast or deep search through the task settings. Also, for different types of artifacts, users should be able to set the desired image matching precision to regulate the accuracy of the monitoring results. Different visual comparison criteria need to be used to identify different types of artifacts. For instance, for 2D objects, such as paintings or book covers, high-precision comparison can be applied for exact matching, which will greatly reduce the probability of false alarms. If necessary, different image features and matching approaches can be applied for different types of 2D and 3D artifacts.

In addition, in order to avoid tracing, some artifacts may not be sold as a whole, but may first be dismantled into several pieces and then each part is sold separately. Or the artifacts may be damaged during theft or transportation and only broken fragments can be sold. This requires that the applied image matching tool not only be able to recognize complete artifacts but also have the ability to identify parts of artifacts.

References

[1] Kimberly Alderman. Honor Amongst Thieves: Organized Crime and the Illicit Antiquities Trade. Indiana Law Review, Forthcoming, vol. 45, no. 3, pp. 602-627. 2012.

[2] Neil Brodie: Virtually gone! The Internet market in antiquities. In Proceedings of the 6th International Conference of Experts on the Return of Cultural Property. Seoul: Overseas Korean Cultural Heritage Foundation, pp. 190–204, 2017.

[3] Neil Brodie: How to Control the Internet Market in Antiquities? The Need for Regulation and Monitoring. Antiquities Coalition Policy Brief No 3, July 2017.

[4] York Yannikos, Marc Leon Agel, Julian Heeger, Simon Bugert: Cooking Spiders: Efficient OSINT with Chefs and Recipes. In Proc. IS&T Int'l. Symp. on Electronic Imaging: Media Watermarking, Security, and Forensics, 2025.

[5] Simon Bugert, Huajian Liu, Waldemar Berchtold, Martin Steinebach: Cultural assets identification using transfer learning. In Proc. IS&T Int'l. Symp. on Electronic Imaging: Imaging and Multimedia Analytics at the Edge, pp. 273-1 - 273-4, 2022.

[6] Alexander Kolesnikov, Lucas Beyer, Xiaohua Zhai, Joan Puigcerver, Jessica Yung, Sylvain Gelly, Neil Houlsby: Big transfer (BiT): General visual representation learning. In Proc. of ECCV, pp. 491-507, 2020.

[7] Catawiki: https://www.catawiki.com/

[8] eBay: https://www.ebay.com/

Author Biography

Huajian Liu received his B.S. and M.S. degrees in electronic engineering from Dalian University of Technology, China, in 1999 and 2002, respectively, and his Ph.D. degree in computer science from Technical University of Darmstadt, Germany, in 2008. He is currently a senior research scientist at Fraunhofer Institute for Secure Information Technology (SIT). His major research interests include information security, digital watermarking, robust hashing and digital forensics.

York Yannikos received his Diplom (equiv. Master's degree) in computer science from the University of Rostock, Germany in 2008. Since then he has worked as research associate in the Media Security and IT Forensics department at the Fraunhofer Institute for Secure Information Technology (SIT) and at the National Research Center for Applied Cybersecurity (ATHENE) in Darmstadt, Germany. His research interests include darknet marketplaces, open source intelligence, and digital forensic tool testing.

Julian Heeger is a research associate in the Media Security and IT Forensics department at the Fraunhofer Institute for Secure Information Technology (SIT) and a researcher at the National Research Center for Applied Cybersecurity (ATHENE) in Darmstadt, Germany. He holds a Master's degree in IT security from the Technical University of Darmstadt.

Simon Bugert received his Master's degree in computer science from the Technical University of Darmstadt, Germany in 2021. Since then, has been a research associate in the Media Security and IT Forensics department at the Fraunhofer Institute for Secure Information Technology (SIT) and at the ATHENE National Research Center for Applied Cybersecurity.

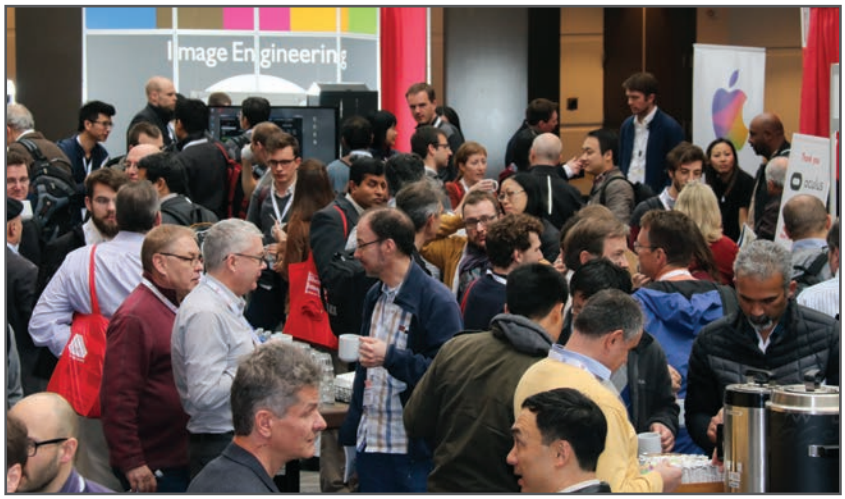
Waldemar Berchtold has headed the Multimedia Security research group since 2022 at the Fraunhofer Institute for Secure Information Technology (SIT) and is a researcher at the National Research Center for Applied Cybersecurity (ATHENE) in Darmstadt, Germany. He received his diploma in mathematics in 2008 and his Ph.D. in 2022 at TU Darmstadt. The focus of his research is in various areas of multimedia security for authenticity and integrity proof and digital watermarking. He has led numerous projects in the field of media security with a focus on audio and video.

Prof. Dr. Martin Steinebach is the manager of the Media Security and IT Forensics division at Fraunhofer SIT. From 2003 to 2007 he was the manager of the Media Security in IT division at Fraunhofer IPSI. He studied computer science at the Technical University of Darmstadt and finished his diploma thesis on copyright protection for digital audio in 1999. In 2003 he received his PhD at the Technical University of Darmstadt for this work on digital audio watermarking. In 2016 he became honorary professor at the TU Darmstadt.

JOIN US AT THE NEXT EI!

electronic IMAGING

Imaging across applications . . . Where industry and academia meet!



• SHORT COURSES • EXHIBITS • DEMONSTRATION SESSION • PLENARY TALKS •
• INTERACTIVE PAPER SESSION • SPECIAL EVENTS • TECHNICAL SESSIONS •

www.electronicimaging.org

ADVANCING IMAGING SCIENCE 75 1947 IS&T 2022 imaging.org