

The Design of a Multimedia-Forensic Analysis Tool (M-FAT)

Shahlaa Mashhadani^{1,3}, Hiba Al-kawaz^{1,4}, Nathan Clarke^{1,5}, Steven Furnell^{1,5,6}, Fudong Li²

¹Centre for Security, Communications and Network Research, Plymouth University, UK

²School of Computing, University of Portsmouth, UK

³Computer Science Department, College of Education for Pure Science, Ibn Al Haytham, Iraq

⁴Computer Science Department, College of Science for Women, Baghdad University, Iraq

⁵Security Research Institute, Edith Cowan University, Perth, Western Australia, Australia

⁶Centre for Research in Information and Cyber Security, Nelson Mandela Metropolitan University, South Africa

Abstract

Digital forensics has become a fundamental requirement for law enforcement due to the growing volume of cyber and computer-assisted crime. Whilst existing commercial tools have traditionally focused upon string-based analyses (e.g., regular expressions, keywords), less effort has been placed towards the development of multimedia-based analyses. Within the research community, more focus has been attributed to the analysis of multimedia content; they tend to focus upon highly specialised specific scenarios such as tattoo identification, number plate recognition, suspect face recognition and manual annotation of images. Given the ever-increasing volume of multimedia content, it is essential that a holistic Multimedia-Forensic Analysis Tool (M-FAT) is developed to extract, index, analyse the recovered images and provide an investigator with an environment with which to ask more abstract and cognitively challenging questions of the data. This paper proposes such a system, focusing upon a combination of object and facial recognition to provide a robust system. This system will enable investigators to perform a variety of forensic analyses that aid in reducing the time, effort and cognitive load being placed on the investigator to identify relevant evidence.

1. Introduction

With the enormous increase in the number of images, videos, and audio recordings available, multimedia evidence has come to play a fundamental role in criminal investigations [1, 2]. The significant increase in the volume of photographs (images*) and video context is having a direct impact on the time and cost of investigations, with much of the current effort resulting in investigators having to manually analyse the context. Consequently, the forensic investigators require a set of forensic analyses to enable them to more efficiently identify relevant evidence [3]. As a result, Forensic Image Analysis (FIA) has emerged as a new branch of digital forensics that enables the investigators to effectively

and accurately extract evidence from a huge number of images in an automatic and forensically sound manner [4]. However, there are at present many challenges that still exist. For instance, forensically, little work has been undertaken using object and face recognition to better understand the context of images. Should an investigator wish to identify all images with a particular object in, they would need to manually investigate each item. Likewise, should an investigator be interested in a particular individual (possibly the suspect) and wish to understand within the sources available, who this individual has interacted with, again, using current tools, a manual inspection and verification would be required. Whilst facial recognition could be utilised, current implementations only operate well within a very constrained set of external conditions (namely front-facial images with a consistent illumination) which often are not present within cases. Existing forensic tools such as EnCase and the Forensic Toolkit (FTK) are insufficient in areas such as automatic content image analysis, extraction of evidence, facial recognition, and in identifying and correlating images [5].

The aim of this paper is to present a novel and holistic multimedia forensic analysis system that can aid the investigation process in analysing, interpreting and correlating multimedia-based context. The proposed automated framework will be able to analyse a large volume of image sources in an efficient and accurate manner to create the necessary annotation and features (AF) that can be utilised to inspect, correlate and analyse the evidence. This will reduce the cognitive burden placed on the investigator when handling large volumes of data and thus provide more timely analysis of the data.

The remainder of the paper is organized as follows. Section 2 provides an overview of the current state of art within object and face recognition. Building upon this, Section 3 is devoted to the proposed M-FAT system architecture and processes that underpin the approach. Section 4 provides an insight of the M-FAT prototype followed by the

illustration of its user interface windows that will demonstrate various functionalities exhibited by the prototype. Section 5 presents a discussion. The conclusion is listed in Section 6.

2. Background

Image recognition can best be analyzed under two methods (i.e., object and the more specialized face recognition) in order to comprehend the current capabilities and limitations. Efforts have been made to narrow down the search environment, so that the investigation is focused to the most current states of the art in object and facial recognition. The research methodology has utilized a range of keywords (object-based image retrieval (centric object retrieval, non-centric object retrieval), and multiple object-based image retrieval, followed by automatic image annotation studies, facial recognition, partial or disguise faces, facial aging, illumination, face pose and expression) to research related studies from various academic databases IEEE, Google Scholar, and Science Direct. The keyword “forensic” is used to find which studies are more related with this field.

Five criteria are applied to select the papers, these are: all publications less than two pages long (including posters, presentations, abstracts, or short theoretical papers) are excluded; non-peer-reviewed publications are eliminated; the language of this literature review is English; site number, impact factor, and publication year.

2.1. Object Recognition

Few studies have focused upon image analysis for the purpose of digital forensics and identifying and extracting evidence from images [6]. An analysis of these studies is summarized in Table 1.

Some of these studies have offered good procedures for FIA and achieved high retrieval accuracy. However, they suffer from the fact that it deals with a specific criminal case. In addition, they have suffered from limitations in their work, such as undetermined number of images that used for experiments or analysis, or they only use a small volume of pictures. In addition, no criteria were applied to evaluate the performance, or no comparison with other studies was performed [7–10]. Moreover, the special characteristics of forensic images are different from characteristics of standard images; therefore, the image features that are suitable to describe standard image databases are inefficient for forensics. For example, the background of forensic photographs is typically far more complicated than those used within the experimental studies, because the target object could be damaged, deficient, or the object may appear very small in the

picture [3]. In addition, the clarity and accuracy of forensic image retrieval are essential requirements for any investigation; however, some real-life images suffer from noise or losing blocks such as losing a number of bits, when sending the image through a wireless channel, and thus require enhancement before analysis [11]. Manual image annotation is yet another challenge, because annotating image manually needs big cost, time consuming, etc. [12].

The findings highlight that there has been little work performed on the subject of extracting evidence from images or solving criminal cases through FIA. Moreover, very few studies are able to overcome the challenges of finding and discovering forensically interesting and suspicious or beneficial patterns within huge datasets while taking into account the requirements of accuracy and speed.

In order to overcome the above problems, research from existing fields such as Object-Based Image Retrieval (OBIR) and Automatic Image Annotation (AIA) could be employed on forensic images to current challenges of image analysis within the forensic domain. However, the forensic examiner needs an automatic system that can recognize multiple objects in the same image, although these objects may differ in size, colour, shape, texture, and orientation. Despite a considerable amount of literature having been published in OBIR, the main limitation is focusing upon having a single main object only. They tend not to focus upon real-life complex imagery. The experiments for these studies were also conducted on only a small and very specific number of images [13–15]. In addition, image retrieval accuracy decreases dramatically with an increasing number of images [16, 17]. Furthermore, there is a substantial gap between low-level content features (color, shape, etc.) that are used for OBIR and semantic concepts (e.g., keyword, text, descriptor) used by humans to interpret images. Moreover, in this approach, users must have an example or a query image at hand, because the query must be an image [18].

As it already mentioned, previous OBIR methods suffer from several kinds of issues. Consequently, AIA systems could be used instead of an OBIR to describe images with words in place of using image features. AIA is a process of automatically assigning words to a given image and it suggests a promising way of achieving more efficient image retrieval and analysis, by bridging the semantic gap between low-level features and high-level semantic contents in image access [19].

This will enable the ability to search based upon keywords and solve problems presented by OBIR systems. Therefore, AIA is considered a highly valuable tool for image search, retrieval, and archival systems [20]. However, AIA studies suffer from

Table 1. Summary of FIA studies

Ref.	Object Extraction Method	Features Extraction	Performance (%) ⁺		Database Name	#Images
			Precision	Recall		
[3]	---	Colour and texture		62 70	forensic and Corel databases	400 forensic images and 800 images
[6]	Background subtraction algorithm	Scale-Invariant Feature Transform (ASIFT) and min-hash technique	85	---	Three videos	203 vehicle object images
[7]	---	Colour, texture, and shape	---	---	---	---
[8]	---	Grey Level Co-Occurrence Matrix (GLCM), texture	---	---	fired bullets, firing pins, extractor marks, ejector marks, and cartridges	50 images
[9]	Region Of Interest (ROI)	Histogram, texture, entropy and Speeded-Up Robust Features (SURF)	98	---	---	250 images
[10]	---	Colour, texture and shape	---	---	---	---
[11]	---	Filtering algorithm and Reconstructing algorithm	median filter	---	---	---
[12]	---	Scale-Invariant Feature Transform (SIFT)	90	--	tattoo images from Michigan State Police	64,000 tattoo images

+ Some results are approximated from studies.

multiple problems such as there is no standard annotation database that has been utilised to evaluate system performance, and most studies conduct experiments using unrealistic image databases [21, 22].

In addition, there is a great disparity in system performance, because of the divergence in segmentation, features, and classifier approaches, as well as the number of images that used in the systems assessment [23, 24].

Furthermore, it should be noted that studies have proposed solutions to the problems of multiple objects retrieval and AIA associated with image retrieval systems, and have achieved high retrieval accuracy. Even then, there is still a problem that none of these studies tested images related to forensic cases and real-life complex and diverse imagery. This makes it impossible to determine whether these studies would achieve high performance in FIA. Moreover, the forensic case images are changeable that makes it difficult to build for each case own AIA system.

2.2. Face Recognition

Face recognition has become more popular in forensics; however, a number of issues within a forensic context still need to be addressed. The efficiency of face recognition is affected by internal and external factors. Internal factors include

uncooperative people in front of the camera such as pose variation, facial expression, faces occluded, accessories and aging [25]. External factors are unrelated to the user, such as light factors, camera quality, and more than one person in the same location, which could obscure the subject's face [26]. In order to improve the efficiency of the forensic facial recognition, these barriers should be thoroughly investigated. Some of the prior facial recognition studies are summarized in Table 2 and then discussed in the paragraphs that follow. Focusing upon facial aging, several studies adopted generative face images according to age progression to minimize the age gap in face matching technique. [27] introduce one example of these studies; they used the craniofacial growth during formative years up to age 18 to improve the recognition accuracy. However, the drawback was ignoring the face texture growth such as fat tissue (that could be an essential feature in the analysis process). While [28] generated series of age-progressed face photos between 1 and 80 years. They dealt with face shape and texture changes, which made the results close to the reality, but they depended on human decisions instead of an automatic identification system.

In comparison, other studies preferred to use the discriminative approach to solve the facial aging issue in face recognition system by using the local features of the face, which they consider is more robust to age variation [29, 30].

Table 2. Summary of facial recognition studies

Ref.	Approach	Recognition Accuracy (%) ⁺⁺	Database (Subjects, Images)
Facial Aging			
[27]	Shape growth modelling	15.0	Private database (109,233)
[28]	Automatically age progression	---	The Google Images (-, 40000)
[29]	Discriminative model	83.9	MORPH album 2 (20569,78207)
		47.5	FG-NET (82,1002)
[30]	Multiview discriminative model	65.2	MORPH album 2 (20569,78,207)
		91.8	FG-NET (82,1002)
Facial pose			
[31]	Mosaicing scheme	96.76	CMU PIE (68, 494)
		97.06	WVU Multispectral (40, -)
[32]	Gabor-based method	86.8	FERET (200, 1196)
		67.6	CMU PIE (68, 494)
[33]	3D transformation model	99	CMU-PIE (68, 494)
		95.6	FERET (200, 1400)
[34]	3D features model	95.31	FERET (200, 1400)
Illumination Factor			
[35]	The maximum filter	98.9	Yale B (10, 5760)
		94.44	extended Yale B (38, -)
[36]	the shadow compensated technique	99	CMU-PIE (68, 494)
		92.3	Yale B (10, 5760)

⁺⁺ Some results are approximated from studies.

Moreover, they combined multi-feature descriptors to obtain more face discriminative information that could support the recognition system. However, the results are vulnerable to other issues such as pose change that could produce the low accuracy [29].

Human interaction in front of surveillance cameras has added new challenge in the forensic system. For example, head pose (e.g., frontal face or not), and partial face or occluded (e.g., face hidden by glasses, hat, and scarf). In spite of numbers of researchers have sought to overcome the facial pose issue in face recognition the limitations are low pose degree (e.g., 20o degree), one direction pose (e.g., horizontal face), and determine individual images required to process the system [31]. In addition, recognition accuracy decrease with an increasing of face pose degree [32]. Other researchers preferred to correct the face pose by creating a 3D face viewing from a 2D

image [33, 34]. In some cases, the 3D model makes the system more robust due to the high discriminative information. Nevertheless, the 3D model needs additional time for processing images database.

Illumination factors also play a key role in the matching processing, holding a significant impact on the overall system performance. A number of studies have attempted to minimize illumination effects on images to increase the recognition accuracy. One of these studies proposed a method of filtering images with illumination variation to obtain smooth images for face recognition [35]. Moreover, [36] proposed a shadow compensated technique that adding weighted average intensity to light angles instead of shadow variations on the facial image. The problems of illumination in previous face recognition studies were due to limitations such as face pose, light angle, the capture environment (indoors, outdoors, night, etc.), and the face image noise and its effect on recognition accuracy.

Regarding multiple image issues, a small number of studies have tried to cope with multiple challenges within the face recognition system [37, 38]. They investigated a face recognition system based on facial expression, face pose, and illumination issues. Their systems only applied and evaluated three issues on non-real-life images individually. However, in their system the database does not include real-life images. On the other hand, [39] proposed a study to identify any suspect person in a large crowd of people with uncontrolled captured images. However, their system focused on partial face images rather than other image issues.

In order to improve the speed of the facial matching system, Park and Jain filtered database by using gender, and ethnicity as demographic information that does not change over time [40].

As demonstrated above, existing studies have attempted to deal with the different effects application of facial recognition. To the best of our knowledge, there have not been attempts to solve all issues together in one system. Additionally, [41] conducted a study of the Boston Marathon bombings of 2013 and analyzed the reasons why the automated face recognition system failed to identify the suspected persons at the time. Their study concluded that forensic facial recognition system operates under unconstrained conditions of people in the presence of digital surveillance cameras. Therefore, the current forensic systems require further investigation to overcome the drawbacks of them.

3. M-Fat Architecture

The objective of the proposed system is about incorporating image analysis within a case management-based system that goes beyond the

current state of the art both within forensics and within their specific specialist domains. The key requirements are:

- Acquire and process a wide variety of base forensic images and live sources (e.g. computer, mobile, cloud, CCTV).
- Describe the photos nature and its contents by creating the necessary annotation and features to recognize object-based and faces.
- Provide a range of forensic analyses and correlation capability to aid an investigator in querying the image source.

In analyzing and creating the necessary AF, the proposed M-FAT will seek to overcome the aforementioned weaknesses of existing annotation and facial recognition systems to provide an effective and robust multimedia forensic analysis tool. The proposed framework is illustrated in Figure 1. This architecture consists of the following processes:

- **M-FAT Manager:** The M-FAT manager represents an interface between the investigator and the underlying processes that help the investigator with managing the overall system. In this stage, the investigator can work with the current cases or create a new case, in addition to manage users and configuration settings. The manager is also responsible for managing authentication, authorization and accountability infrastructure to ensure only the relevant investigators get access to cases they are assigned and to ensure integrity and chain of custody are maintained.
- **Evidence selection:** this stage involves collecting all videos and images from different sources such as CCTV camera, mobile, digital camera, computer images, hard drive and manual data. The system will refine the collection data through the exclusion of irrelevant images based on image metadata as identified by the investigator, in order to facilitate the process of selecting the target image. The output from evidence selection goes to image annotation and face feature extraction processes in order to recognize evidences and save the results are stored in the process evidence database, or to M-FAT manages to select the requested evidence.
- **Image annotation:** This stage explains the process for extracting and annotating objects (evidences) for each image. It consists of the following processes:
 - **Privacy process:** Sometimes an image includes a label or text in its content, such as a name, a car registration number, or a personal address, which may be considered as private information. Thus, this process will reveal whether the image includes any

private information and, if so, the image will be saved in a separate list in order to address it individually. The images that are saved in the separate list will be tackled separately by hiding important information through the use of a mask and then send it to external AIA systems, or by sending it to one secure external AIA system. If there is no private information, the image will be sent to the next stage directly.

- **Multiple AIA systems:** In this process, the image will be sent to multiple external existing AIA systems such as CLOUD VISION API and Microsoft Cognitive Services (Computer Vision API) for annotation. Then, the results from these systems will be collected and sent to the next stage. Multiple AIA systems will be employed in order to annotate these objects, and then their outputs are fused in order to improve the accuracy of annotation results over the results that can be achieved through employment of a single automatic annotation system. The AIA systems will describe the image with words that name the object; however, these systems are unable to describe features of the object such as color or shape. Therefore, the proposed system will improve upon the annotation process by adding an annotation of these features to the external AIA system results.

- **Fusion process:** This process will be utilized to fuse the results from multiple AIA systems to provide more efficient results than the individual systems involved in the process.

- **Face Feature Extraction:** This stage determines faces from image and extracts their features. A detailed of each process will explain below:

- **Face Detection:** In the first process, the facial area will be determined and extracted from the image. In order to get the best performance outcome, a number of techniques will be used to seek best accuracy. After cropping the face from the entire image, the face will be normalized; and various face components such as eyes, mouth, and nose can be located.

- **Feature Extraction:** In this process, a multi feature extraction systems will be utilized to extract effective face features that can be used to improve face recognition quality (each one could to focus on different face features). The results of each system will save as vectors then sent them to the next process.

- **Fusion:** The fusion engine will build to fuse face features from multiple feature vectors. This process will be tried to improve the overall accuracy by increasing the dimension of feature space. The final features will consider as evidences and save in the process evidence database.

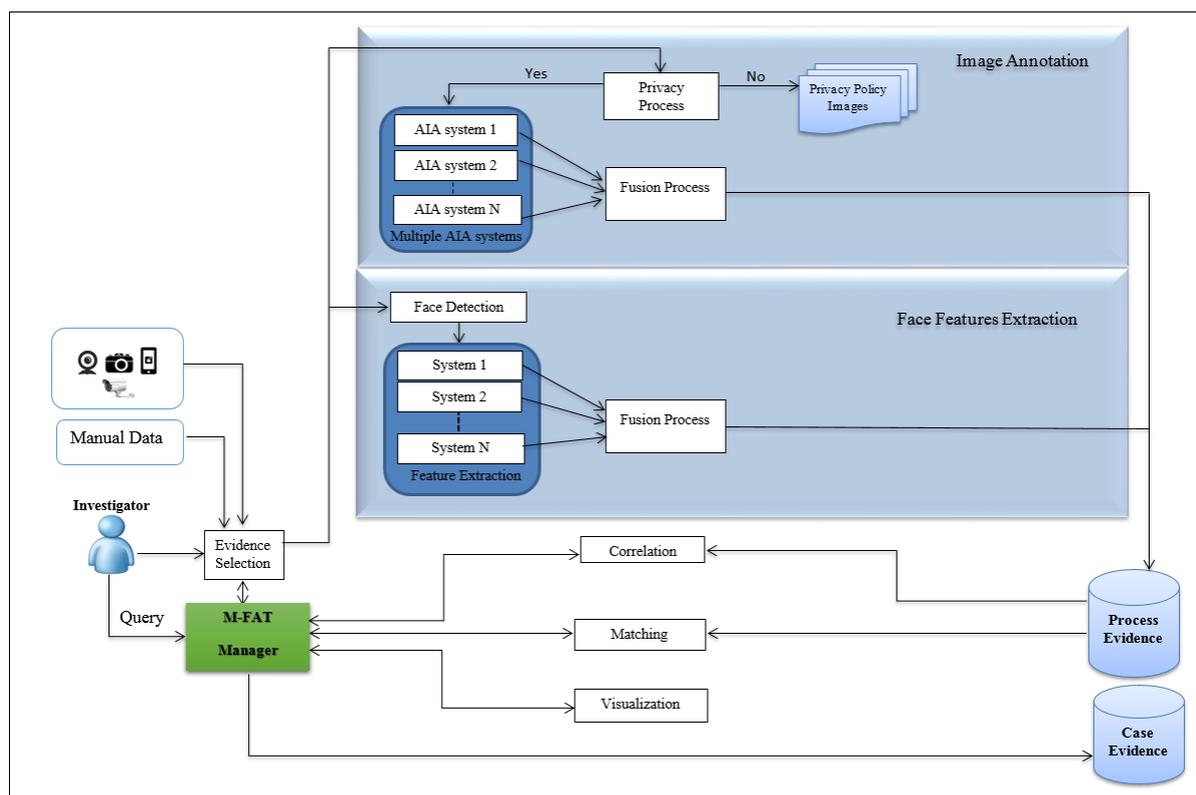


Figure 1. M-FAT architecture

- **Matching:** This stage includes establishing a search engine connected to the process evidence database, which has the capability of accommodating single, or multiple keywords or query image. After that, the system will retrieve all images that satisfy the search conditions. For example, if the investigator inserts the text 'red car' with requested face image, then the system will retrieve all images that contain all conditions.
- **Correlation:** The system will correlate the retrieved images through finding the relationships that connect between them by using multiple approaches. The aim of correlation process is to enable the investigator to ask higher-level and more abstract questions of the data, thereby construct the crime scene and find the relatives evidences. In addition, rather than looking through hundreds, possible thousands of images, the investigator would be given a smaller number of images of the specific content and metadata.
- **Visualization:** Data visualization enables the investigator to see analytics presented visually and assisting them to better understand complex concepts. For instance, google map, graph and report will be used to present the results.

4. M-Fat Prototype

This prototype aims to reflect how the M-FAT prototype would help digital investigators to make the search process less daunting and time-consuming, and it will also improve the search results by finding answers to the essential questions in the investigation: what, who, why, how, when, and where. Therefore, it will assist the investigator to find relevant pieces of evidence from among others. Once the investigator has logged in, the system will automatically direct him to the dashboard.

The dashboard (as shown in Figure 2) represents an interface between the investigator and the underlying processes that help him with managing the overall system, which includes three options: New/ Edit Case, Global Sitting and Account Management. In the first one the investigator can edit current cases information, and also open any case in order to add new evidence for selected case, show case history, carry out advance analytics, visualize results that could be used as a potential evidence. Adding new case is the other function that provides by this interface, all information that related to the case are fed to the system by the investigator, such as case number, case name, case date, investigators' names and associated information. After adding the case, the investigator will be able to add image (copy) that related with the case from various sources such as forensic image, smart phone, CCTV or database. The

system implements many preprocessing for acquired data such as hashing, converting videos to images, extracting metadata, filtering, image resizing and enhancement. The system provide the ability to acquire data (photos) from multiple sources for each case, these data will be sent to other processes in order to be annotate and extract face features that will be saved in order to use it later. Moreover, the investigator will also be able to archive the case, when it is closed or when there is no need to work on it.

Global setting undertakes the system settings that contain preferences that apply to all system parts. These settings include user configuration, configure archiving, specify automatic image annotation and face recognition systems that will be used.

In the last option in dashboard, the administrator has the ability to add user's information, and also set the privileges that permit to the user to perform his tasks. In order to ensure that access to each part of the system is performed by authorised users, different roles are defined.

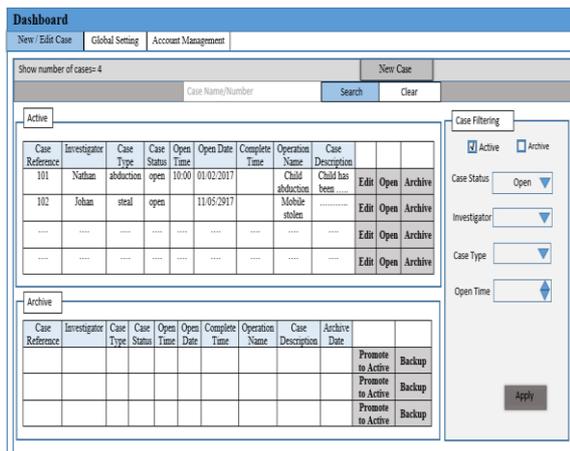


Figure 2. M-FAT dashboard

Once the case is created and their resources are acquired and examined, the system provides the investigator with case dashboard as illustrated in Figure 3. The search choice connects with the Process Evidence database, which has images, annotations, face features and metadata. The function of the search process is to find the similarity between images based on text query, which include single, or multiple words, and/or image query, then refine the results based on chosen filters. The system will use the text and/or image query to search the Process Evidence database to find all images that contain text query within their annotations and similar the query image, and also meet all selected filters. The system will able to candidate the retrieved results based on a combination of multiple filters, in addition to confidence score value, number of retrieval images

that need to be reviewed, and also the relation between words (And/Or).

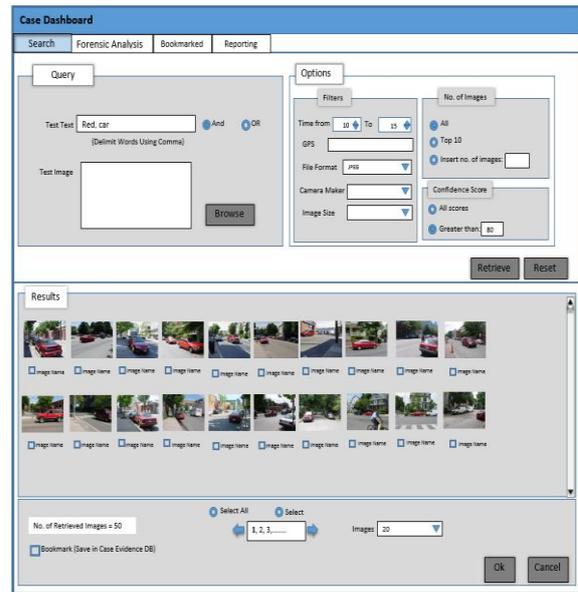


Figure 3. Case dashboard

Moreover, the analysis will be carried out depending on previous-selected images or new-selected images. After the selection process is performed, the system will introduce multiple options that include different forensic analysis processes. These include various types of image comparison approaches that match between images depending on face features, text, GPS information and metadata. The objective from using different approaches is to enable the investigator to correlate between relevant images based on which analysis would be most appropriate for the type of requested evidence in order to reduce the search domain, find the requested evidence in a short time and show the relationship between images to draw complete picture of the crime. The forensic analyses choice will provide multiple options such as metadata filtering, object matching, GPS tracking and social network analysis. Information is shared between screens based on user role.

In the first option, the investigator will refine the interesting images from the retrieval results (search choice), using metadata provides useful information that can help the investigators to determine the exact location of a photo that was captured or obtain information about the device owner from the model or the serial number that collected in the photo's metadata. Therefore, the forensic investigator can track down suspects through excluding all irrelevant images based on image metadata, which identified by the investigator, in order to facilitate the process of selecting the target images.

In object matching, the search results will be refined through image content matching. The object recognition will be used to find, from a query image

(selected image/images), identical or similar images in the search result or database. The similarity between images depend on object recognition, shape or color, this means that it depends on upon the content of an image, rather than on textual information. Once the similarity comparison has been done, all related images will be queried and retrieved and displayed to the investigator depending on the matching level. For instance, in search screen the investigator could retrieve all red cars, however, the number of retrieved images is large. Consequently, the investigator could select the suspect car from retrieved images and use object matching to find all images that exactly contain the same selected car.

From a forensic point of view, the location data (possibly from GPS coordinates) are valuable, because it gives an overview of the last locations of a suspect or provides an accurate movement profile of a person. Therefore, the geographical tracking utilizes the Google Maps that will provide an overview of what directions a person/an object was utilized, and thereby specify their whereabouts. This can be helpful in solving criminal cases such as kidnappings and runaway youths to drug trafficking and homicides. As a result, the basic purpose of a geographical tracking system is to track a suspect, target vehicle or other objects through locating and viewing the images on Google Map based on GPS information.

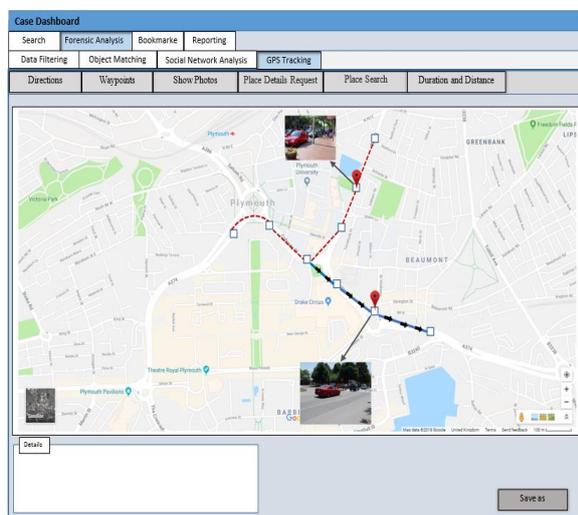


Figure 4. Geographical tracking

Where multiple paths are possible, the system will provide a probabilistic measure indicating which to investigate first as shown in Figure 4.

Social network analysis is looking for people face through facial matching and return all photos for the suspect person. In addition, it has ability to return all person faces that appear together with suspect person in the same photo. This forensic analyzing uses a

graph theory to represent a network structure which is characterize as nodes (people faces), and edges or links which is represented the matching results as shown in Figure 5. Furthermore, filtering options will be provided to aid the investigator depending on investigation requirements. For example, it could display only the suspect images that found in databases or display all faces images which their appearance associate with the suspect.

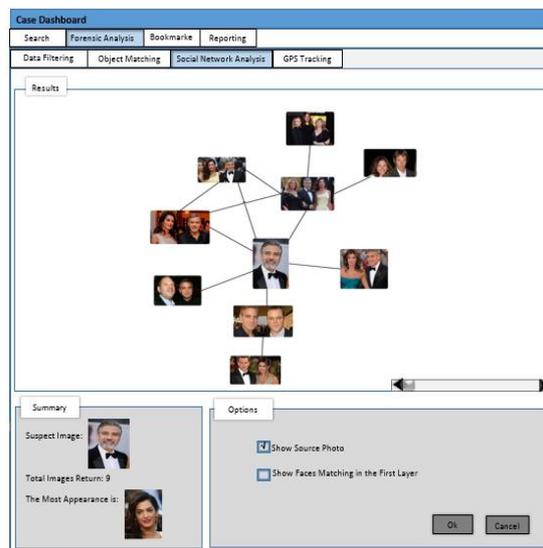


Figure 5: Social network analysis

Bookmark option is utilized to store both search results and correlation results. All searching and correlating results and their related information such as timestamp of search and query value are stored in the case evidence database in order to come back to it for more analysis or printing report.

Finally, reporting is utilized to create a report that will be prepared for submission to the court, or for returning to the results later when required. The work that has been performed in all previous processes is documented and presented during the reporting stage. The investigator could create the final report that contains the requested results, in addition to all case information.

5. Discussion

In reviewing the literature, no study has combined object and face in order to describe image contents. In addition, most studies have some drawbacks in their works especially that they have not been tested on real-life complex imagery and whenever the number of images is increased, their retrieval accuracy will decrease dramatically. Furthermore, there is no study that deals with all facial image issues together in one system. As a result, the premise of the proposed M-FAT is to combine between object and face in one system that meets

investigator requirements, and capitalize upon existing research and systems in a multi-algorithmic manner to benefit from the different feature extraction and classification approaches. In order to achieve the proposed system requirements: Firstly, various techniques such as FTK and Mobile Phone Examiner Plus (MPE+) will be used to acquire images from different sources (e.g. CCTV, computer and mobile) based on the source type. In addition, the acquired images that need to be investigated, suggesting that these images are usually large in number, vary in quality, unconstrained illumination, various orientation, object size, irregular background, and contain multiple objects. As a result, these images are voluminous and need processing often in near real-time and in doing so maintain the level of accuracy. The quality of images is another an important issue in image analysis, because the reliability of any inspection task is based on that quality. Therefore, the image under consideration should be checked first to determine whether the image quality is sufficient to allow for a meaningful and reliable analysis

Secondly, different multiple object and facial recognition systems that have the ability to recognize different objects and faces with different characteristics from the image will be examined and their results will be fused, in order to improve the evidence extraction process. The objective of using multiple systems is to overcome limitations of each system individually and looking for different reliable information. Previous researches within biometrics in particular have shown this to be beneficial [42, 43]. However, further work needs to explore the extent to which a multi-systems approach would work in this specific application.

Finally, accuracy and speed of retrieving images are the biggest challenges facing the use of image analysis in digital forensics. However, once annotated or query image, merely looking at all the results of a single or set of keywords or image will not necessarily diminish the investigative task. Therefore, the M-FAT system will tackle this challenge by applying additional knowledge to the retrieved images with the aim of enabling the investigator to ask and filter evidence using a wider range of information. Therefore, the prototype is designed to use the multi-algorithmic approach that collects different annotations and face features for the same image that gives the investigator the ability to use many different words and suspect/victim face picture to describe the whole image content in order to increase the accuracy of annotation and face recognition to get advantages of all of them. By employing various image analysis techniques for correlating between images based on the type of evidence, the framework is capable of finding all images that have the same evidence. The investigator is able to select the type of analysis and the database

that used for comparison based on crime requirements. The framework implements the recursive operation in order to decrease search domain and find evidence in a short time, in addition, the framework stores all search and correlation results to utilize it later in another investigation. Furthermore, multiple visualize forms are used to view the results in order to show the relatives results. Regarding to operational considerations that should be taken for the proposed system: big storage is essential in order to save images and their features (AF and metadata) in a database, addition to the retrieval results that also should be saved to review it again in any time. Therefore, multiple tables are utilized in the proposed system to store all information and results, because the variation in type of information that needs to be stored, in addition to usage multiple levels of analysis. In addition, the use of publically available AIA or facial recognition systems results in benefitting from the latest developments of image analysis, without having to develop and manage the system, it does introduce the problem of submitting evidence to an external untrusted source for analysis. Consequently, pre-processing procedures are introduced the necessary privacy that required.

6. Conclusions

Current forensic tools provide a basic level of analysis for multimedia-based content. With increased volumes and heterogeneity of data to process and the timeliness of analysis often being key issues, specific tools need to be developed. The M-FAT prototype is able to deal with various images contents that collected from a different source by utilizing a combination of image content analysis techniques that permitting of obtaining results that are more accurate. The M-FAT is designed to use the multi-algorithmic approach via fusion to achieve the best results that will permit investigators to retrieve multiple pieces of evidence from a large and heterogeneous forensic image database efficiency. By employing various image analysis techniques for correlating between images based on the type of evidence, the prototype is capable of finding all images that have the same evidence. The investigator is able to select the type of analysis that used for comparison based on crime requirements. Furthermore, multiple visualize forms are used to view the results in order to show the relatives results. Current stage of the research has managed to finalize the M-FAT prototype in order to exam the system efficacy to deal with forensic cases.

Future research should, therefore, concentrate on the development of the prototype and evaluate the performance and effectiveness of the M-FAT

towards verifying the ability to meet all the key requirements.

7. References

- [1] A. K. Jain, B. Klare, and U. Park, "Face Matching and Retrieval in Forensics Applications," pp. 20–28, 2012.
- [2] A. Singh, "Exploring Forensic Video And Image Analysis." 2015.
- [3] H. Yuan and L. Ying, "Study on forensic image retrieval," in 2014 9th IEEE Conference on Industrial Electronics and Applications, 2014, no. 112, pp. 89–94.
- [4] R. B. Hanji and V. Rajpurohit, "Forensic Image Analysis - A Frame work," *Int. J. Forensic Comput. Sci.*, vol. 8, no. 1, pp. 13–19, 2013.
- [5] M. Al Fahdi, N. L. Clarke, F. Li, and S. M. Furnell, "A suspect-oriented intelligent and automated computer forensic analysis," *Digit. Investig.*, vol. 18, pp. 65–76, 2016.
- [6] Chao-Yung Hsu, Li-Wei Kang, and H.-Y. Mark Liao, "Cross-camera vehicle tracking via affine invariant object matching for video forensics applications," in 2013 IEEE International Conference on Multimedia and Expo (ICME), 2013, pp. 1–6.
- [7] C. Wen, D. Ph, and C. Yu, "Image Retrieval of Digital Crime Scene Images," pp. 37–45, 2005.
- [8] R. S. Choraś, "Texture Based Firearm Striations Analysis for Forensics Image Retrieval," in *Advances in Intelligent Systems and Computing*, vol. 184 AISC, 2013, pp. 25–31.
- [9] K. V Shriram, P. L. K. Priyadarsini, and A. Baskar, "An intelligent system of content-based image retrieval for crime investigation," vol. 7, pp. 264–279, 2015.
- [10] S. A. Gulhane and A. A. Gurjar, "Content based Image Retrieval from Forensic Image Databases," vol. 5, no. 3, pp. 66–70, 2015.
- [11] A. Aljarf and S. Amin, "Filtering and Reconstruction System for Gray Forensic Images," vol. 9, no. 1, pp. 20–25, 2015.
- [12] J. Lee, W. Tong, R. Jin, and A. K. Jain, "Image Retrieval in Forensics: Application to Tattoo Image Database," *IEEE Multimed.*, 2011.
- [13] J. Wu, X. Wang, and H. Xing, "Regional objects-based image retrieval," in 2011 Chinese Control and Decision Conference (CCDC), 2011, pp. 1273–1277.
- [14] M. Mohammadpour and S. Mozaffari, "A method for Content-Based Image Retrieval using visual attention model," in 2015 7th Conference on Information and Knowledge Technology (IKT), 2015, pp. 1–5.
- [15] M. Shamsujjoha, M. S. Ahmed, F. Hossain, and T. Jabid, "Semantic modelling of unshaped object: An efficient approach in content-based image retrieval," in 2014 17th International Conference on Computer and Information Technology (ICCIT), 2014, pp. 30–34.
- [16] N. Gupta, S. Das, and S. Chakraborti, "Revealing What to Extract from Where, for Object-Centric Content Based Image Retrieval (CBIR)," in *Proceedings of the 2014 Indian Conference on Computer Vision Graphics and Image Processing - ICVGIP '14*, 2014, pp. 1–8.
- [17] N. W. U. D. Chathurani, S. Geva, V. Chandran, and V. Cynthujah, "Content-Based Image (object) Retrieval with Rotational Invariant Bag-of-Visual Words representation," in 2015 IEEE 10th International Conference on Industrial and Information Systems (ICIIS), 2015, pp. 152–157.
- [18] D. Zhang, M. Monirul Islam, and G. Lu, "Structural image retrieval using automatic image annotation and region based inverted file," *J. Vis. Commun. Image Represent.*, vol. 24, no. 7, pp. 1087–1098, Oct. 2013.
- [19] C. Jin and S.-W. Jin, "Automatic image annotation using feature selection based on improving quantum particle swarm optimization," *Signal Processing*, vol. 109, pp. 172–181, Apr. 2015.
- [20] S. H. Amiri and M. Jamzad, "Automatic image annotation using semi-supervised generative modeling," *Pattern Recognit.*, vol. 48, no. 1, pp. 174–188, 2015.
- [21] Y.-F. Huang and H.-Y. Lu, "Automatic Image Annotation Using Multi-object Identification," in 2010 Fourth Pacific-Rim Symposium on Image and Video Technology, 2010, pp. 386–392.
- [22] M. Hidajat, "Annotation Based Image Retrieval using GMM and Spatial Related Object Approaches," vol. 8, no. 8, pp. 399–408, 2015.
- [23] Y. Xia, Y. Wu, and J. Feng, "Cross-Media Retrieval using Probabilistic Model of Automatic Image Annotation," vol. 8, no. 4, pp. 145–154, 2015.

- [24] J. Majidpour, E. Khezri, H. Hassanzade, and K. S. Mohammed, "Interactive tool to improve the automatic image annotation using MPEG-7 and multi-class SVM," in 2015 7th Conference on Information and Knowledge Technology (IKT), 2015, pp. 1–7.
- [25] S. Z. Li and A. K. Jain, *Handbook of Face Recognition*. 2010.
- [26] X. Xu, W. Liu, and L. Li, "Low Resolution Face Recognition in Surveillance Systems," vol. 2014, no. January, pp. 70–77, 2014.
- [27] N. Ramanathan and R. Chellappa, "Modeling Age Progression in Young Faces," 2006.
- [28] I. Kemelmacher-shlizerman, S. Suwajanakorn, and S. M. Seitz, "Illumination-Aware Age Progression," 2014.
- [29] Z. Li, U. Park, and A. K. Jain, "A Discriminative Model for Age Invariant Face Recognition," vol. 6, no. 3, pp. 1028–1037, 2011.
- [30] D. Sungatullina, J. Lu, G. Wang, and P. Moulin, "Multiview Discriminative Learning for Age-Invariant Face Recognition," 2013.
- [31] R. Singh, S. Member, M. Vatsa, and S. Member, "A Mosaicing Scheme for Pose-Invariant Face Recognition," vol. 37, no. 5, pp. 1212–1225, 2007.
- [32] L. A. Cament, F. J. Galdames, K. W. Bowyer, and C. A. Perez, "Face recognition under pose variation with local Gabor features enhanced by Active Shape and Statistical Models," vol. 48, pp. 3371–3384, 2015.
- [33] A. Asthana, T. K. Marks, M. J. Jones, K. H. Tieu, and R. Mv, "Fully Automatic Pose-Invariant Face Recognition via 3D Pose Normalization," 2011.
- [34] D. Yi, Z. Lei, and S. Z. Li, "Towards Pose Robust Face Recognition," pp. 3539–3545, 2013.
- [35] A. Nabatchian and M. Ahmadi, "An Efficient Method for Face Recognition under Illumination Variations," pp. 432–435, 2010.
- [36] S. Choi, C. Choi, and N. Kwak, "Face recognition based on 2D images under illumination and pose variations," vol. 32, pp. 561–571, 2011.
- [37] F. Bhat and M. A. Wani, "Elastic Bunch Graph Matching Based Face Recognition Under Varying Lighting, Pose, and Expression Conditions," vol. 1, no. 8, pp. 51–59, 2015.
- [38] M. Sultana, M. Gavrilova, and S. Yanushkevich, "Expression, Pose, and Illumination Invariant Face Recognition using Lower Order Pseudo Zernike Moments," 2014.
- [39] S. Liao, A. K. Jain, and S. Z. Li, "Partial Face Recognition: Alignment-Free Approach," vol. 35, no. 5, pp. 1193–1205, 2013.
- [40] U. Park and A. K. Jain, "Face Matching and Retrieval Using Soft Biometrics," vol. 5, no. 3, pp. 406–415, 2010.
- [41] J. C. Klontz, E. Lansing, and E. Lansing, "A Case Study on Unconstrained Facial Recognition Using the Boston Marathon Bombings Suspects The Boston Marathon Bombings - Investigation Timeline," pp. 1–8, 2013.
- [42] K. V Awalkar, S. G. Kanade, and D. V Jadhav, "A Multi-modal and Multi-algorithmic Biometric System Combining Iris and Face," 2015.
- [43] G. Sathish, "Multi-algorithmic IRIS Recognition," vol. 38, no. 11, pp. 13–21, 2012.

8. Acknowledgment

This project is part of a PhD research currently being carried out at Centre for Security, Communications and Network Research (CSCAN), Plymouth University, U.K. The deepest gratitude and thanks to Baghdad University, Ministry Of Higher Education & Scientific Research and Higher Committee for Education Development in Iraq (HCED) for funding this PhD research.