A Novel Multimedia-Forensic Analysis Tool (M-FAT)

Shahlaa Mashhadani[1]
Centre for Security, Communications and Network Research
Plymouth University
Plymouth, UK
shahlaa.mashhadani @ plymouth.ac.uk

Hiba Al-kawaz[2]
Centre for Security, Communications and Network Research
Plymouth University
Plymouth, UK
hiba.al-kawaz @ plymouth.ac.uk

Nathan Clarke[3]
Centre for Security, Communications and Network Research
Plymouth University
Plymouth, UK
N.Clarke @ plymouth.ac.uk

Steven Furnell[3,4]
Centre for Security, Communications and Network Research
Plymouth University
Plymouth, UK S.Furnell @ plymouth.ac.uk

Fudong Li
Centre for Security, Communications and Network Research
Plymouth University
Plymouth, UK
fudong.li @ plymouth.ac.uk

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.**

Keywords- Multimedia, Forensic Image Analysis, Annotation, Face Recognition.

I. 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

---

[1] Computer Science Department, Collage of Education for Pure Science, Ibn Al Haytham, Baghdad, Iraq.
[2] Computer Science Department, College of Science for Women, Baghdad University, Baghdad, Iraq.
[3] Security Research Institute, Edith Cowan University, Perth, Western Australia, Australia.
[4] Centre for Research in Information and Cyber Security, Nelson Mandela Metropolitan University, Port Elizabeth, South Africa.
* In this research, the term image refers to a picture, photograph or other file that is typically associated with file extensions such as BMP, JPEG, and TIFF.

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 II provides an overview of the current state of art within object and face recognition. Building upon this, Section III is devoted to the proposed M-FAT system architecture and processes that underpin the approach. Section IV provides an illustrative case study that highlights the advantages of the proposed approach. Section V presents a discussion. The conclusion is listed in Section VI.

II. BACKGROUND

Image recognition can best be analysed under two methods (i.e., object and the more specialised 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.

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 I.

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 was 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 retrieve specific evidence and thus to solve many of the current challenges of image analysis within the forensic domain. However, the forensic examiner needs an automatic system that can recognise 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 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].

TABLE I.   SUMMARY OF FIA STUDIES

Ref.	Object Extraction Method	Features Extraction	Performance (%) [+]		Database Name	#Images
			Precision	Recall		
[3]	---	Colour and texture		62	forensic and Corel databases	400 forensic images and 800 images
				70		
[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.

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 a high performance in FIA. Moreover, the forensic case images are changeable that makes it difficult to build for each case own AIA system.

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 camera such as pose variation, facial expression, faces occluded, accessories and aging [25]. External factors are unrelated with 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 II 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 important 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 solving 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]. 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., 20° 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 need additional time for processing images database.

TABLE II. 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.

Illumination factors also play a key role in the matching processing, holding a significant impact upon the overall system performance. A number of studies have attempted to minimize illumination effects on images in order 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 face 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 identity 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 analysed the reasons why the automated face recognition system failed to identity 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 in order to overcome the drawbacks of them.

III. M-FAT ARCHITECTURE

The objective of the proposed system is about incorporating image analysis within a single 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).

- To analyse and create the necessary AF (object-based or facial) to describe the nature of the image.

- To provide a range of forensic analyses and correlation capability to aid an investigator in querying the image source.

In analysing 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:

• Evidence selection: The first stage of the proposed system involves the forensic investigator collecting all videos and images from different source 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 in order to recognize evidences and save the results in process evidence database, or to M-FAT manages to select the requested evidence.
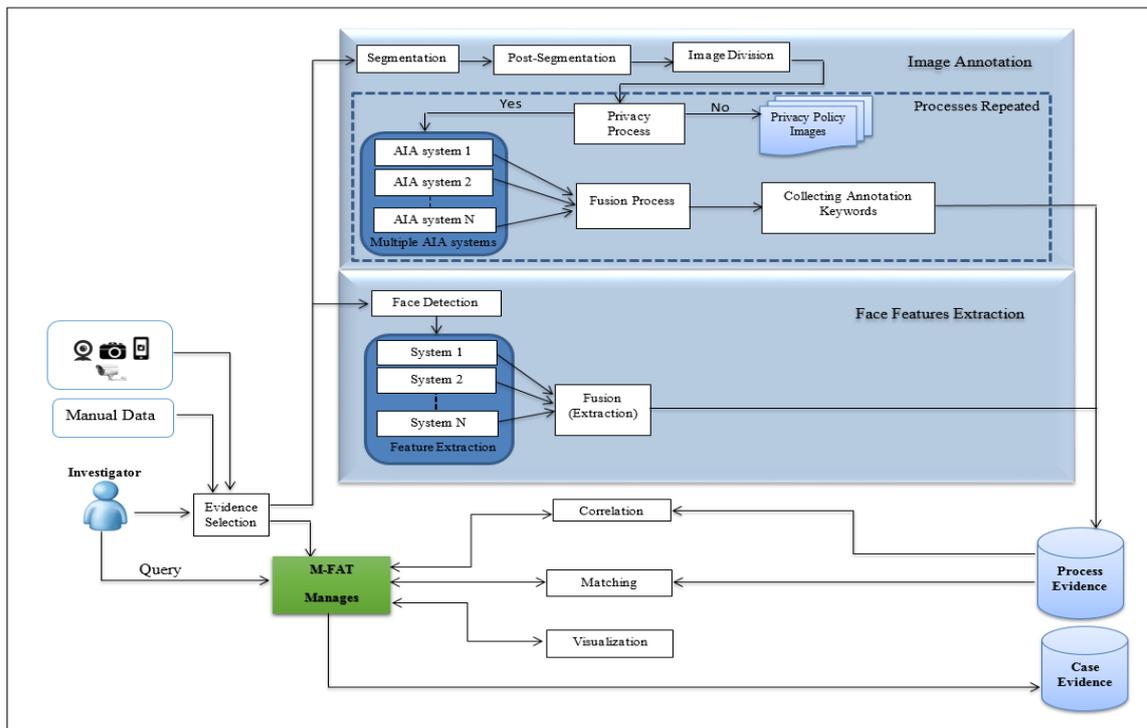
FIGURE 1. M-FAT ARCHITECTURE.

• Image annotation: This stage explains the process for extracting and annotating objects (evidences) for each image. It consists of the following processes:

- Segmentation: An image segmentation process will be carried out in order to divide the image into regions that permit a more meaningful and easier analysis. Forensic images include multiple objects with complex backgrounds; therefore, more than one algorithm will be used in order to extract multiple objects efficiently and effectively.

- Post-segmentation: Various problems that affect the accuracy of object extraction may appear after the segmentation process. These problems include noise, bumpy boundaries, unlinked boundary segments, objects with common boundaries, weak segments, over-segmentation, and edge segments that do not belong to realistic boundaries. It is thus very important for the post-segmentation process to resolve these problems, because the accuracy of extracting objects (evidence) will be inefficient without this step.

- Image division: In this process, the image will be divided into small pieces; and each piece contains one main object, making it ready for the annotation process. Each piece will be called an Object Image (OI). The main goals of this stage are to increase the annotation accuracy and to maintain image privacy before sending it to multiple external AIA systems.

- Privacy process: Sometimes an OI 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 important information. Thus, this process will reveal whether the OI 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 OI will be sent to the next stage directly.

- Multiple AIA systems: In this process, the OI 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 OI 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 utilised to fuse the results from multiple AIA systems to provide more efficient results than the individual systems involved in the process. Combining annotation results from different systems will result in performance improvement.

- Collecting annotation keywords: The result from the previous process will be saved along with the image in order to build the final image annotation. The privacy process, multiple AIA systems, and fusion process stages are repeated depending on the number of objects extracted from the forensic image. Then, the final result (image with their annotated keywords) will be saved in a process evidence database.

• **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 utilised 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 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.

• **M-FAT Manages**: The M-FAT manages is an interface between the investigator and the underlying system to provide the ability to search, correlate and visualize the data. The results from this interface will be saved in the case evidence database. Based on the requirements, the investigator can recall the aforementioned results that could be used as a potential evidence.

• **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**: A Decision Support System (DSS) will be used at this stage to facilitate the role of the investigator by finding correlations between retrieval images based on metadata and AF, in order to construct the crime scene. This process will assist the investigator to find relevant pieces of evidence from among others.

• **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.

IV. CASE STUDY

To help illustrate how the proposed framework would operate, a child abduction example is presented. In this example, it is assumed that a child has been kidnaped. Intelligence provides a rough last location for the child and information that they were seen getting into a red car. In order to solve a child abduction case, an investigator starts to collect all preliminary evidence that may help to find the child as fast as possible. For example, narrowing the timeframe of abduction, examining properties of the car that a witness believes was involved in the abduction, determining the location of the abduction, and any information about suspect (e.g., face description, age and gender). The next step of the investigative process would involve collecting all available imagery (e.g. videos from surveillance cameras at the crime scene and from nearby surveillance systems). Manual analysis of the sources in and around the timeframe would provide an investigator with an image of the child's face and of the car she was forced into. Timely analysis and evidence and the reporting of the investigators findings is critical to the safe recovery of the child.

The current solution would involve teams of investigators manually trawling through the footage from possibly dozens of evidence sources. The use of a manual human matching process is a laborious and time-consuming resulting in examining large volumes of image data and given the pressurised nature of the task likely to result in a high proportion of human error.

The proposed system will permit an investigator to select the necessary evidence sources, automatically process all of the footage. The investigator will be able to select the objects of interest – in this case the face of the child and the car that she got into. The system will then perform facial and object recognition across the evidence sources, providing an investigator with a prioritised set of results with which to interact with. The system will refine the retrieval results based on metadata (time, location, and date of the abduction) in order to reduce the number of retrieval results. The investigator will be able to target image (the suspect's car) from the retrieval results, and the DSS will provide further correlation and analysis functions that would enable the target car or face to be tracked across the different evidence sources. The resulting visualisation would provide the graphical map of the resulting journey alongside the image sources utilised to identify the path of the car. Where multiple paths are possible, the system will provide a probabilistic measure indicating which to investigate first as shown in Figure 2.
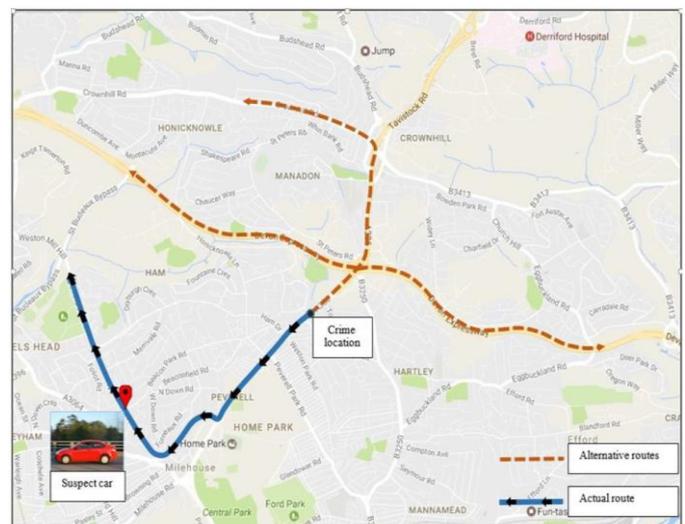

Figure 2. Example of car tracking result.

V. DISCUSSION

In reviewing the literature, no study was combined between 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 not 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 meet investigator requirements, and capitalise 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 FTK and Mobile Phone Examiner Plus (MPE+) will be used to acquire images from different sources (e.g. CCTV, computer and mobile) base 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 recognise 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 and 44]. 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 proposed 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, it is important to develop the DSS that can link the annotation and image feature alongside relevant metadata to enable investigators to ask higher-level more abstract questions of the data. Consequently, more investigation will be required to find the best correlation methods to provide the necessary functionality that required building a relation between various types. Moreover, additional work required to find the best way to present the results in different models based on investigator requirements such as symbol map, node-link diagram, and chart.

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. 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. Therefore, pre-processing procedures should be introduce the necessary privacy are required.

VI. CONCLUSIONS

Current forensic tools provide a basic level of analysis for multimedia-based content. With increased volumes of data to process and the timeliness of analysis often being key issues, specific tools need to be developed. The paper proposed a holistic multimedia architecture using a multi-algorithmic approach to enhance the power of final recognition result. Current stage of the research has managed to utilise multiple systems and fuse their results to recognise image contents and has identified that compared to the individual systems the fusion results are more promising. This has shown positive signs in terms of the feasibility of the proposed-system-achievement making the research goal attainable. Future paper(s) will discuss the completed results.

Future work however needs to investigate how the proposed system will be able to support the investigator in providing complex and high-level questioning of the resulting forensic data.

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.

[44] N. AI-Naffakh, N. Clarke, P. Dowland and F. Li, "Activity Recognition using Wearable Computing", in Proceedings of the 11th International Conference for Internet Technology and Secured Transactions (ICITST-2016), Barcelona, 2016, pp. 189-195.