Advanced Facial Recognition for Digital Forensics

Hiba Al-kawaz[1,2], Nathan Clarke[1,3], Steven Furnell[1,3,4], Fudong Li[1, 5], Abdulrahman Alruban[1,6]
[1]Centre for Security, Communications and Network Research, Plymouth University, Plymouth, UK.
[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 University, Port Elizabeth, South Africa.
[5]School of Computing, University of Portsmouth, Portsmouth, UK.
[6]Computer Sciences and Information Technology College, Majmaah University, Majmaah, Saudi Arabia.
hiba.al-kawaz @ Plymouth.ac.uk
N.Clarke @ Plymouth.ac.uk
S.Furnell @ Plymouth.ac.uk
Fudong.li @ port.ac.uk
Abdulrahman.alruban @ Plymouth.ac.uk

Abstract:: Forensic facial recognition has become an essential requirement in criminal investigations due to the advent of electronic devices such as CCTV, digital cameras, mobile phones, and computers and the huge volume of content that exists. Forensic facial recognition goes beyond facial recognition in that it deals with facial images under unconstraint and non-ideal conditions, such as poor image resolution, facial orientation, illumination, expression, and the presence of accessories. These conditions have a huge impact on the recognition performance. A wide variety of facial recognition algorithms exist, each more or less susceptible to various environmental conditions. This paper proposes a multi-algorithmic fusion approach by utilising multiple commercial facial recognition systems to overcome particular weaknesses in singular approaches to obtain the best facial identification accuracy. The advantage of focusing upon commercial systems is because it releases the forensic team from developing and managing their own solutions and subsequently also benefits from state of the art updates in underlying recognition performance. A set of experiments were conducted to evaluate three commercial facial recognition systems (Neurotechnology, Microsoft, and Amazon Rekognition) to determine their individual performance using facial images with varied conditions. The second experiment sought to determine the benefits of fusion. Two challenging facial datasets were identified for the evaluation; the first was a publically available dataset known as 'CAS-PEAL-R1'. The second dataset represents a more challenging yet realistic set of digital forensics scenarios collected from publically available celebrity photographs. The experimental results have proved that using the developed fusion approach achieves better identification rate as the best tested commercial system has scored 67.23%, while the multi-algorithmic fusion system scored an accuracy of 71.6%.

KEYWORDS
Forensic, Face Recognition, Fusion, Multi-algorithm

1 Introduction

In recent years, facial recognition systems have received a great deal of attention in digital forensics, particularly because of the enormous volume of digital photographs and video databases that can be used for identifying individuals (Jain et al., 2012). Moreover, as information technology has evolved and spread worldwide, the use of facial images in passport systems and by police forces further augmented facial image databases. In 2014, about 245 million video surveillance systems were installed around the world (Jenkins, 2015). In the UK alone, there were between 4 million to 5.9 million CCTV cameras and the USA deployed approximately 30 million surveillance cameras (Jain et al., 2012). These CCTV systems create billions of hours footage on a weekly basis, offering an enormous amount of information that can be used to track suspects when crimes occur (Jain et al., 2012). By utilising facial recognition technologies, valuable information used in the detection of culprits can be extracted from photos or videos that are found at crime scenes (Peacock et al. 2004). As a result, automating the process of suspect recognition can save forensic investigators' immense volumes of time in comparison to if the search task was carried out manually by watching each video in the database.

During a digital forensic investigation, all available evidence is collected and later carefully analysed. The analysis process categorizes files into different types (e.g., images and videos) then an investigator can view and examine these files separately. The use of photos and video data in an investigation helps investigators to track suspects faces, their locations, time, persons appeared with them, and their activities (Carrier, 2003). However, uncooperative subjects, low quality of pictures and poor illumination make the recognition process a challenging task for any digital forensic facial recognition system. For example, for the Belgium's Zaventem Airport attack in 2016, limitations of the forensic face recognition system had a serious impact on the ability to perform successful recognition, delaying the identification of the suspects in a time-critical investigation (Shoichet et al., 2016). Since the system failed to track a suspect because of the distance between the camera and person resulting in a low resolution of the faces and variation in illumination. Furthermore, different expression and face poses can also cause a failure in facial recognition.

As demonstrated by the examples above, visual evidence and facial recognition, in particular, are valuable investigative tools. Nevertheless, although a large amount of academic and commercial research effort is given to the area, automatic facial recognition still suffers from several significant drawbacks in the achievement of accuracy, delaying in tracking suspects or failing to identify the suspect. Despite, a range of commercial software focussed upon face recognition having being developed, most existing forensic tools suffer from the inclusion of these identification techniques within their analysis. For example, Forensic Toollkit (FTK) is a good tool for analysing most files but lack in any facial identification capacity. As a result, forensic facial recognition has been identified as a key objective of the FBI's next-generation program (FBI, 2015).

A large number of facial datasets have been collected and publicly available due to the importance of facial recognition in different emerging areas (e.g. computer vision, security, and digital forensics), including the FERET dataset (NIST, 1993), Multi-PIE dataset (Gross et al., 2010), and MORPH dataset (Ricanek and Tesafaye, 2006). However, most of these datasets are not compatible with forensic investigations as the images were taken in a constrained or semi-constrained fashion. They are not indicative of real-world imagery. This paper aims to evaluate three widely known commercial facial recognition algorithms using a realistic facial dataset containing celebrities face photos to simulate the real challenge for forensic cases by exploring the capability to recognize uncooperative faces. In addition, the study will explore the application of fusion to create a multi-algorithmic approach to determine whether performance can be improved.

The rest of the paper is organized as follows. Section 2 highlights the related work in the area of facial recognition. Section 3 explains the experimental methodology. Section 4 presents the experimental results. Section 5 discusses the findings and their impact; and, the paper concludes with Section 6.

2 Related Work

Recently, facial recognition has become popular in forensic investigations; however, a number of issues within the forensic context need to be considered before the technology is fully utilised. The efficiency of face recognition is directly affected by many issues which include uncooperative people in front of the camera, face pose, expression, accessories and varied age (Li and Jain, 2011). Whilst other issues are related to the lighting and camera resolution (Xu et al., 2014). Most studies in the facial recognition field have focused on tackling a single issue only. For example, several studies have adopted generative face images according to age progression to minimize the age gap in face matching (Kemelmacher-Shlizerman et al., 2014; Ramanathan and Chellappa, 2006). While other studies preferred to use the discriminative approach by using the local features of the face (Li et al., 2011; Sungatullina et al., 2013). In addition, studies were interested in detecting and processing head pose as it is considered a primary factor that affects recognition accuracy (Singh et al., 2007; Cament et al., 2015). One of the common approaches to correct the face pose is by creating a 3D face viewing from a 2D image (Asthana et al., 2011; Yi et al., 2013). Also, Nabatchian et al. (2010) and Choi et al. (2011) both proposed a method to minimize illumination effects on images in order to increase the recognition accuracy.

In comparison, a limited number of studies have tried to manage with multiple challenges within the facial recognition systems. Bhat and Wani (2015) and Sultana et al. (2014) investigated a face recognition system based on face expression, face pose, and illumination issues. However, their systems only applied and evaluated three issues individually on non-real life images. Liao et al. (2013) proposed a study to identify any suspicious person in a large crowd of people with uncontrolled captured images. However, their system mainly focussed on partial face images rather than other image issues.

Therefore, forensic systems require further investigation in order to overcome the drawbacks of existing studies and identifying unconstrained face images. Klontz and Jain (2013) conducted a study of the Boston Marathon bombings of 2013 and analysed the reasons why the automated face recognition system failed to identify the suspected persons at that time. They used three commercial matchers with high accuracy. Their study concluded that forensic facial recognition systems operate under unconstrained faces of people in the presence of digital surveillance cameras and need more progress to overcome such issues. Moreover, Wang et al. (2017) used a couple of commercial off the shelf (COTS) systems to search for persons within large scale photos by using deep features. Whilst, Best-Rowden et al. (2016) evaluated one of COTS face matchers on their Newborns, Infants, and Toddlers Longitudinal face image database to explore the ability of face identification on children faces. Their result showed that facial recognition technology still have complexity to recognize young children faces. In addition, Juefei-Xu et al. (2015) studied the performance of COTS face recognition systems on partial faces or occluded facial parts. However, they found inconsistencies existed in the COTS systems depending on images sources especially occlusion faces.

As demonstrated above, existing studies have attempted to deal with the different effects application of facial recognition further considerable progress in commercial facial recognition technique. However, the real forensic scenarios need more effort in the facial recognition tools that have a significant impact in the criminal investigation field.

3 Experimental Methodology

The aim of this experiment to evaluate the performance of three commercial facial identification systems (Neurotechnology, Microsoft, and Amazon Rekognition) on different facial image issues. This is followed by an experiment seeking to improve performance through a holistic system using fusion. Three experiments were developed to investigate this achieve the aim of this study as follows:

- The first experiment is a controlled experiment to determine the performance of the commercial face recognition algorithms when applying a standard facial dataset (CAS-PEAL-R1) with varying facial issues such as illumination, accessories, expression, and pose.
- The second experiment evaluates the examined systems on a carefully collected facial images dataset that simulates real digital forensic scenarios.
- The third experiment seeks to evaluate whether a multi-algorithmic approach to classification would improve underlying classification performance.

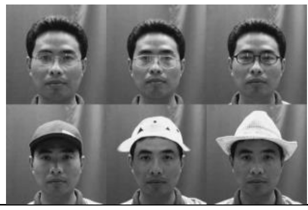
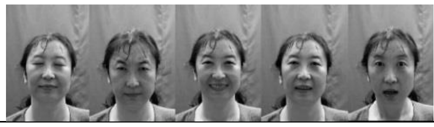
The control experiment provides a basis for understanding how well these commercial systems perform by using standardised photography from publically available facial corpuses. The results of this experiment can then be directly compared to experiment two which focusses upon using more forensically realistic images – where numerous facial recognition challenges are likely to co-exist simultaneously. An analysis of facial recognition research identified several differing routes to classification, with different algorithms focussed upon differing aspects of the facial image. It was this analysis that gave rise to the question of whether a multi-algorithmic fused approach to classification might improve upon the results (i.e. the strengths of one algorithm overcoming the weaknesses of others) and thus the third experiment.

The study focussed upon the use and evaluation of commercial algorithms because in practice this could bring several advantages. It would relieve forensic investigators from having to design, implement and manage facial recognition systems. Leaving the specialist to manage the application-independent facial recognition system will also lead to algorithmic improvements and updates beyond the capability of the forensic team.

3.1 Experiment 1: A Controlled Experiment (CAS-PEAL-R1)

For the first experiment the CAS-PEAL-R1 Chinese face dataset was utilised as it simulates several facial recognition challenges (i.e., pose, expression, lighting, and accessories) (Gao et al., 2008). It was collected at the Chinese Academy of Sciences (CAS) between August 2002 and April 2003. It consists of 30,900 images across 1,040 subjects (595 men and 445 women). In this experiment, only those that meet all conditions (pose, illumination, expression, and accessories) were included (95 subjects); while the remaining were excluded. Detailed information of CAS-PEAL-R1 is illustrated in Table 1, including a breakdown of the dataset with examples to illustrate the nature of the imagery. The evaluation was based upon Identification Rate (IR) for rank-1 (1: Many classification); further False Positive Identification Rate (FPIR) and Failure to Acquire (FTA) rates have also been implemented for further analysis of the data. The FPIR is the ratio of test samples that are classified as true while they are false. Whilst, the FTA is represented rate of failure to create face templates in testing dataset (Accessory, Expression, Lighting, and Pose).

Table 1: Subsets of CAS-PEAL-R1 dataset used in experiment 1.

Subset	No. of Subjects	Images per subject	Total Images	Samples
Normal (enrolment set)	95	1	95	
Accessory	95	6	569	
Expression	95	5	475	
Lighting	95	>=9	1,203	
Pose	95	20	1,900	

3.2 Experiment 2: A Realistic Dataset

In order to fulfil the requirements of the second experiment, a more realistic dataset is required. Although a number of realistic face datasets have been collected from the Internet such as LFW (Huang et al., 2007) and CelebFace (Sun et al., 2013), they do not include enough samples for the subjects under all challenge types (e.g. lighting variation and accessories) that this study tries to tackle. As a result, the facial dataset for this experiment was collected by the authors from the web based upon celebrities as these images could be easily collected. The criteria used in choosing the images was depend on an unconstrained facial in a different environment (day and night), far from camera and close, wearing accessories (glasses and hat), with a different period time. In the end, a total of 4,001 images were initially collected from 100 subjects and each subject has at least 30 images. In addition, 100 frontal images (one image per subject) were collected and used as the reference image (enrolment dataset) while all others images (4,001 images) were used as testing images (testing dataset). Figure

1 illustrates some samples of Adam Sandler images were used in this study. Moreover the same evaluation metrics in experiment 1 are used for this experiment.

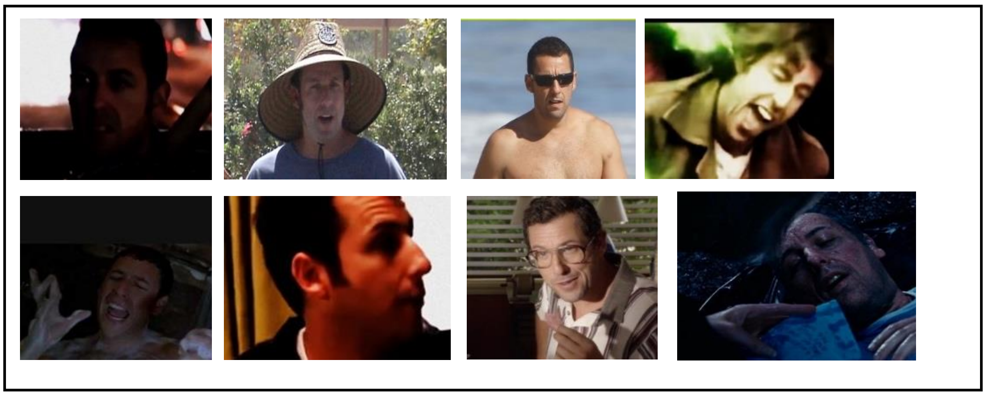


Figure 1: Some samples of Adam Sandler images.

3.3 Experiment 3: The Fusion Search

Prior biometric research has shown that multi-algorithmic fusion has resulted in improved performance, it seemed prudent to explore this aspect (Mitra et al., 2016; Ross and Jain, 2003). The multi-algorithms fusion leverages the knowledge of multiple systems (e.g. Neurotechnology, Microsoft, and Amazon Rekognition identification systems) as illustrated in Figure 2. Normalisation is a crucial characteristic of any multi-algorithmic fusion approach. It can be used to normalize various scores that have been obtained from different matching algorithm. Before the fusion process, these scores are mapped from multiple domains into the public domain. In the fusion approach, the multi-level scores that are generated from multiple face recognition algorithms are combined in order to achieve the single score and to show best performance. Depending on the rank-1 results of each system, weights are added to each of the score levels. The weight value is obtained by studying the accuracy of each system in previous experiments, which leads to having the best score result.
The proposed fusion technique was applied on two datasets that are employed in experiments 1 and 2. The purpose of considering same datasets in this experiment to compare the performance between fusion results and those achieved in experiments 1 and 2.

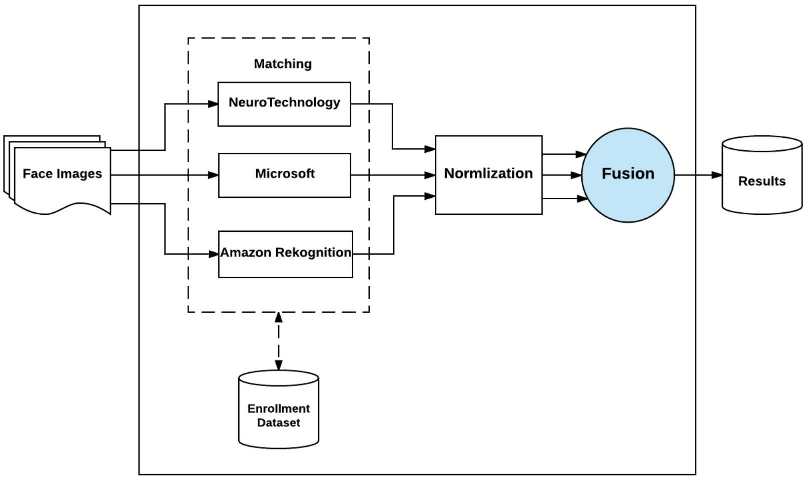


Figure 2: Fusion technique architecture.

4 Experiments Results

4.1 Experiment 1 Results

The result of experiment 1 is illustrated in Table 2. For the accessories category, Microsoft and Amazon achieved the same IR score (at Rank 1) of 98.76%; while their Neurotechnology counterpart got slightly lower result i.e.

92.61%. The IRs of the expression condition part of three systems are all over 98%, with 98.31%, 99.57% and 99.78% for Neurotechnology, Microsoft and Amazon respectively. Regarding the lighting and poses condition, the overall performance of three algorithms decreased, with IRs of 63.42%, 86.69% and 83.95% for Neurotechnology, Microsoft and Amazon accordingly. While Amazon achieved the highest IR of 85.73% regarding the pose issue; it should be noted that Neurotechnology performance dropped significantly to 31.31% for the same category. This could be caused by due to the big face pose angle problem in some samples and limitation in these systems to manage it.

Table 2: Experiment 1 results (the three commercial systems performance for CAS-PEAL-R1 dataset).

Subset	IR at Rank1 (%)		
	Neurotechnology	Microsoft	Amazon Rekognition
Accessories	**92.61**	**98.76**	**98.76**
Expression	**98.31**	**99.57**	**99.78**
Lighting	**63.42**	**86.69**	**83.95**
Pose	**31.31**	**74.47**	**85.73**

Table 3 demonstrates the FPIR and FTA rates of the first experiment. At a glance, the FPIR and FTA rates of all systems for the lighting and pose conditions are significantly higher than those presented in the accessories and expression conditions. Also incorrect matching rates (i.e. FPIR) of three systems in all face conditions are less than FTA rates. This high FTA rates can be caused by some images in testing subset as they were acquired with bad resolution, darkness, bad exposure, and angle pose. This means the templates that are implemented in the matching process will be less than total input images; hence, the high failure in IR rate. For example, Neurotechnology system achieved highest FTA rate 66.1% in pose issue. This means only 34% of face templates were implemented in this system.

Table 3: The FPIR and FTA rates of experiment 1 for CAS-PEAL-R1 dataset.

Subset	Neurotechnology		Microsoft		Amazon Rekognition	
	FPIR (%)	FTA (%)	FPIR (%)	FTA (%)	FPIR (%)	FTA (%)
Accessories	0.52	6.85	0.52	0.70	1.05	0.17
Expression	0.21	1.47	0	0.42	0	0.21
Lighting	11.13	25.27	2.66	10.64	5.90	10.14
Pose	2.05	66.10	0.78	24.73	4.26	10

Overall, the results of this experiment demonstrate the ability of the three systems in facial identification for four facial images issues (accessories, expression, lighting, and pose). It is clear that all the examined commercial systems accuracy suffers when the darkness of the image is increased as well as the facial pose degree.

4.2 Experiment 2 Results

As demonstrated in Table 4, with an IR at Rank 1 of 67.23%, the Microsoft system achieved the best performance against the Celebrities dataset; while the system of Neurotechnology obtained the lowest performance with just 6.6%. In comparison to results obtained from the first experiment, it evidenced the performance of all three systems has dropped significantly due to complexity of realistic facial images of the Celebrities dataset, highlighting the challenge that the digital forensic investigator has to face when dealing with real life scenarios.

Table 4: Experiment 2 results (the three commercial systems performance for Celebrities dataset).

Enrolment subjects	IR at Rank1 (%)		
	Neurotechnology	Microsoft	Amazon Rekognition
100	6.60	67.23	48.24

Table 5 presents the incorrect matching and FTA rates for three algorithms. The highest FPIR of 11.47% was obtained by the Amazon system; while the highest FTA rate of 87.35% was achieved by the Nuerotechnology system. Overall, the FTA rates obtained in this experiment are significantly higher than those are presented in the first experiment. As mentioned previously, the increased FTA rate will decrease the number of face templates that can be sent to next stage (i.e. the matching process). Indeed, the overall results of this set of experiment demonstrate that how realistic photos from the celebrities' dataset affects the perform of top commercial face recognition algorithms, showing the complexity of the problem this paper tries to solve.

Table 5: The FPIR and FTA rates of experiment 2 for Celebrities dataset.

Enrolment subjects	Neurotechnology		Microsoft		Amazon Rekognition	
	FPIR (%)	FTA (%)	FPIR (%)	FTA (%)	FPIR (%)	FTA (%)
100	6.04	87.35	6.37	26.39	11.47	40.28

4.3 Experiment 3 Results

As expected, the identification accuracy obtained by the fusion system have improved in comparison to experiments 1 and 2 results. As illustrate in Figure 3, when applying the proposed fusion technique upon the CAS-PEAL-R1 dataset, improved perform can be observed for all four chosen types; particularly the identification accuracy for the lighting improves from 86.69% by using the Microsoft facial identification system alone to 90.44% (i.e. an improvement of 3.75%). This confirms that the proposed fusion technique outperforms the three individual facial identification systems.

Figure 4 shows the result comparison between the proposed fusion systems and the three chose facial identification systems on the Celebrities data set. Again, the proposed fusion system achieved the best performance of 71.6% of IR; this result is over 4% better than the best individual system (i.e. Microsoft facial identification system) obtained. As a result, this experiment shows that improvement on performance can be obtained by using the proposed fusion system, particularly for the unconstrained dataset; this highlights the potential impact the proposal fusion system may have upon the forensic investigation field.

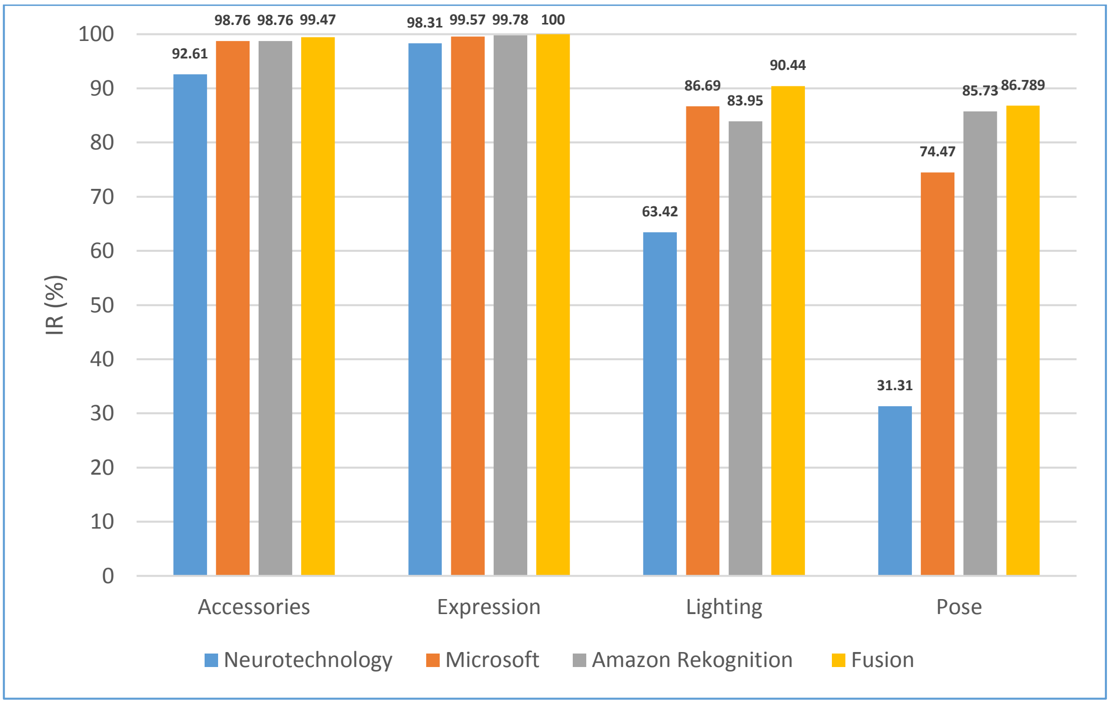


Figure 3: The performance comparison between fusion method and other systems by using CAS-PEAL-R1 dataset.

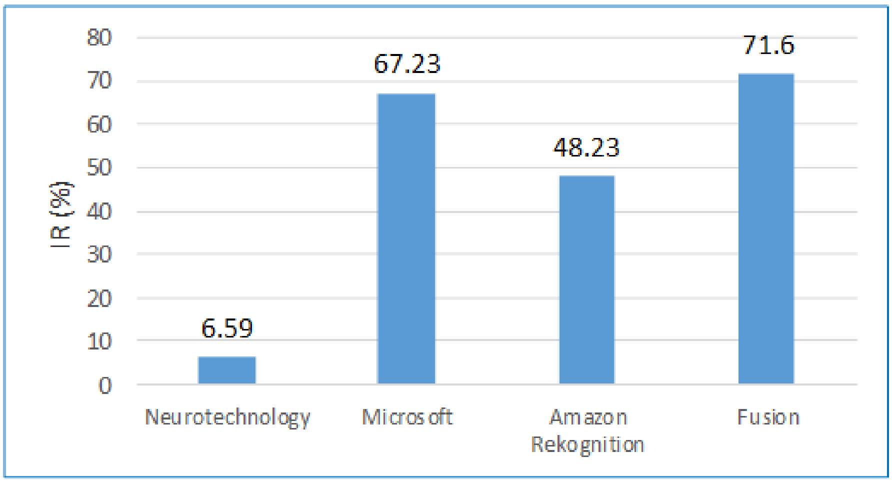


Figure 4: The performance comparison between fusion method and other systems by using Celebrities dataset.

5 Discussion

The observation from the first experiments shows that there is a contrast in performance of the three commercial facial recognition systems (Neurotechnology, Microsoft, Amazon Rekognition) by using CAS-PEAL-R1 dataset. In this experiment, the results indicates how accurate these systems are in matching front face position with hat, glasses, and some expression. The reason for this is that all face samples testing were front with and good resolution. Whilst three systems managed lower facial matching rates with regards to lighting and the lowest in facial pose. The main reason that caused this drop in systems performance is the failure to acquire some templates from input face images due to low image quality, low lighting, and high pose angle. As the template generation is considered the primary step in the facial recognition systems, this led to impact on the number of acceptance templates that are sent to the matching process. Overall, the Microsoft system performed better than the other systems in lighting while Amazon achieved significant performance in pose while the three systems were as same good accuracy in expression and accessories issues. This give a clear impression for each system in front of four face recognition issues.

The second experiment concludes that there is a big drop in facial matching accuracy for three systems comparison with experiment 1. This drop is due to nature of dataset images that was collected for this experiment which were simulate the real world (Celebrities dataset). Moreover, the FPIR slightly increased in this dataset over CAS-PEAL-R1. While there are significantly increased in the FTA rates for all systems when comparing with first experiment. This substantiates that unconstrained facial imagery are a significant challenge in comparison to standard datasets and that commercial systems still struggle with achieving a reliable performance.

As in previous experiments, the performance of the system varies and it is not stable for each system under different image issues. Therefore, the hypotheses of multi-algorithmic fusion approach supports to achieve better performance for the holistic system. The fusion method improves all results of controlled experiment (experiment 1) up 100% in expression issue. In addition, this approach improves facial matching rate at rank-1 from about 67% in Microsoft system (which was best performance) into 71% in realistic dataset. The last enhancement consider an important effect in forensic filed.

This study supports the proposition that a multi-algorithmic approach to forensic facial recognition would improve upon the existing state of the art. Whilst the use of commercial systems have several advantages, more notably a degree of specialisation that should see performance rates maximised, there is an issue of privacy. For example, Microsoft uses cloud services in their recognition process; however, in doing so it saves a copy of the submitted images so that subsequent algorithmic improvements can be made. From a forensic data privacy perspective, this would be a significant barrier to adoption. Notable, other systems, such as Amazon and Neurotechnology do not do this, so it is far from being a standardised approach.

6 Conclusion and Future work

This paper introduced three experiments to evaluate the performance of current commercial facial recognition systems (Neurotechnology, Microsoft, and Amazon Rekognition) and how the multi-algorithmic fusion could improve the accuracy. The multi-algorithm fusion approach showed high accuracy in terms of using face recognition on two selected datasets CAS-PEAL-R1 and Celebrities. When using unconstrained face dataset, the proposed system improves the facial identification accuracy from the highest identification accuracy for commercial systems. Through over the results, the fusion method results has outperformed the other systems. Future work will focus upon exploring how to enhance images to increase recognition performance. For example, using image enhancement as a pre-processing stage prior to classification to re-configure the image for use within existing recognition systems.

References

ASTHANA, A., MARKS, T. K., JONES, M. J., TIEU, K. H. & ROHITH, M. Fully automatic pose-invariant face recognition via 3D pose normalization. Computer Vision (ICCV), 2011 IEEE International Conference on, 2011. IEEE, 937-944.

BEST-ROWDEN, L., HOOLE, Y. & JAIN, A. Automatic Face Recognition of Newborns, Infants, and Toddlers: A Longitudinal Evaluation. 2016 International Conference of the Biometrics Special Interest Group (BIOSIG), 21-23 Sept. 2016 2016. 1-8.

BHAT, F. & WANI, M. A. 2015. Elastic Bunch Graph Matching Based Face Recognition Under Varying Lighting, Pose, and Expression Conditions. *Elastic,* 1.

CAMENT, L. A., GALDAMES, F. J., BOWYER, K. W. & PEREZ, C. A. 2015. Face recognition under pose variation with local Gabor features enhanced by Active Shape and Statistical Models. *Pattern Recognition,* 48**,** 3371-3384.

CARRIER, B. 2003. Defining digital forensic examination and analysis tools using abstraction layers. *International Journal of digital evidence,* 1**,** 1-12.

CHOI, S.-I., CHOI, C.-H. & KWAK, N. 2011. Face recognition based on 2D images under illumination and pose variations. *Pattern Recognition Letters,* 32**,** 561-571.

FBI. 2015. *NEXT GENERATION IDENTIFICATION* [Online]. the FBI's Criminal Justice Information Services (CJIS). Available: https://www.fbi.gov/services/cjis/fingerprints-and-other-biometrics/ngi [Accessed 8-8-2016 2016].

GAO, W., CAO, B., SHAN, S., CHEN, X., ZHOU, D., ZHANG, X. & ZHAO, D. 2008. The CAS-PEAL large-scale Chinese face database and baseline evaluations. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans,* 38**,** 149-161.

GROSS, R., MATTHEWS, I., COHN, J., KANADE, T. & BAKER, S. 2010. Multi-pie. *Image and Vision Computing,* 28**,** 807-813.

HUANG, G. B., RAMESH, M., BERG, T. & LEARNED-MILLER, E. 2007. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical Report 07-49, University of Massachusetts, Amherst.

JAIN, A. K., KLARE, B. & PARK, U. 2012. Face matching and retrieval in forensics applications. *IEEE multimedia,* 19**,** 20.

JENKINS, N. 2015. *245 million video surveillance cameras installed globally in 2014* [Online]. IHS Technology. Available: https://technology.ihs.com/532501/245-million-video-surveillance-cameras-installed-globally-in-2014 [Accessed 27/07/2016 2016].

JUEFEI-XU, F., PAL, D. K., SINGH, K. & SAVVIDES, M. A preliminary investigation on the sensitivity of COTS face recognition systems to forensic analyst-style face processing for occlusions. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2015. 25-33.

KEMELMACHER-SHLIZERMAN, I., SUWAJANAKORN, S. & SEITZ, S. M. Illumination-aware age progression. Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on, 2014. IEEE, 3334-3341.

KLONTZ, J. C. & JAIN, A. K. 2013. A case study on unconstrained facial recognition using the boston marathon bombings suspects. *Michigan State University, Tech. Rep,* 119**,** 120.

LI, S. Z. & JAIN, A. 2011. *Handbook of Face Recognition*, Springer Science & Business Media.

LI, Z., PARK, U. & JAIN, A. K. 2011. A discriminative model for age invariant face recognition. *Information Forensics and Security, IEEE Transactions on,* 6**,** 1028-1037.

LIAO, S., JAIN, A. K. & LI, S. Z. 2013. Partial face recognition: Alignment-free approach. *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* 35**,** 1193-1205.

MITRA, S., WEN, B. & GOFMAN, M. 2016. Overview of Biometric Authentication. *Biometrics in a Data Driven World: Trends, Technologies, and Challenges***,** 1.

NABATCHIAN, A., ABDEL-RAHEEM, E. & AHMADI, M. An efficient method for face recognition under illumination variations. High Performance Computing and Simulation (HPCS), 2010 International Conference on, 2010. IEEE, 432-435.

NIST. 1993. *Face Recognition Technology (FERET)* [Online]. NIST. Available: https://www.nist.gov/programs-projects/face-recognition-technology-feret [Accessed 2015].

RAMANATHAN, N. & CHELLAPPA, R. Modeling age progression in young faces. Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on, 2006. IEEE, 387-394.

RICANEK, K. & TESAFAYE, T. Morph: A longitudinal image database of normal adult age-progression. Automatic Face and Gesture Recognition, 2006. FGR 2006. 7th International Conference on, 2006. IEEE, 341-345.

ROSS, A. & JAIN, A. 2003. Information fusion in biometrics. *Pattern recognition letters,* 24**,** 2115-2125.

SHOICHET, C., BOTELHO, G. & BERLINGER, J. 2016. *Brothers ID'd as suicide bombers in Belgium, 1 suspect 'on the run'* [Online]. CNN. Available: http://edition.cnn.com/2016/03/23/europe/brussels-investigation/ [Accessed 26/07/2016 2016].

SINGH, R., VATSA, M., ROSS, A. & NOORE, A. 2007. A mosaicing scheme for pose-invariant face recognition. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on,* 37**,** 1212-1225.

SULTANA, M., GAVRILOVA, M. & YANUSHKEVICH, S. Expression, pose, and illumination invariant face recognition using lower order pseudo Zernike moments. Computer Vision Theory and Applications (VISAPP), 2014 International Conference on, 2014. IEEE, 216-221.

SUN, Y., WANG, X. & TANG, X. Hybrid deep learning for face verification. Proceedings of the IEEE International Conference on Computer Vision, 2013. 1489-1496.

SUNGATULLINA, D., LU, J., WANG, G. & MOULIN, P. Multiview discriminative learning for age-invariant face recognition. Automatic Face and Gesture Recognition (FG), 2013 10th IEEE International Conference and Workshops on, 2013. IEEE, 1-6.

WANG, D., OTTO, C. & JAIN, A. K. 2017. Face search at scale. *IEEE transactions on pattern analysis and machine intelligence,* 39**,** 1122-1136.

XU, X., LIU, W. & LI, L. 2014. Low Resolution Face Recognition in Surveillance Systems. *Journal of Computer and Communications,* 2**,** 70.

YI, D., LEI, Z. & LI, S. Z. Towards pose robust face recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2013. 3539-3545.