Continuous multibiometric authentication for online exam with machine learning

Short paper

Riseul Ryu
School of Technology, Environment and Design
University of Tasmania
Tasmania, Australia
Email: Riseul.Ryu@utas.edu.au

Soonja Yeom
School of Technology, Environment and Design
University of Tasmania
Tasmania, Australia
Email: Soonja.Yeom@utas.edu.au

Soo-Hyung Kim
Department of Computer Science
Chonnam National University
Gwangju, South Korea
Email: shkim@chonnam.ac.kr

Abstract

Multibiometric authentication has been received great attention over the past decades with the growing demand of a robust authentication system. Continuous authentication system verifies a user continuously once a person is login in order to prevent intruders from the impersonation. In this study, we propose a continuous multibiometric authentication system for the identification of the person during online exam using two modalities, face recognition and keystrokes. Each modality is separately processed to generate matching scores, and the fusion method is performed at the score level to improve the accuracy. The EigenFace and support vector machine (SVM) approach are applied to the facial recognition and keystrokes dynamic accordingly. The matching score calculated from each modality is combined using the classification by the decision tree with the weighted sum after the score is split into three zones of interest.

Keywords Continuous authentication, multibiometric, facial recognition, keystrokes dynamics, score level fusion

1 Introduction

Conducting academic assessment using online exams or internet-based exams has become of great interest to educators because of the reduced time, cost and human errors that the technology offers. There are many attempts to build trust-worthy and secure environments for online exams to ensure academic integrity; for example, single-factor authentication using usernames and passwords. However, online exams still face the challenges in terms of preventing and detecting breaches of academic integrity through cheating and identity falsification (Asep and Bandung 2019).

Authentication is the first step toward building trust-worthy and secure environments which defend against the impersonation of authorised users (Velásquez et al. 2018). As single factor authentication suffers from significant pitfalls (Dasgupta et al. 2016), continuous authentication using multiple factors is an ongoing trend to provide secure, resilient and robust user verification repeatedly throughout online sessions (Traore et al. 2017).

The aim of the research is to design and implement a continuous multimodal biometric system of combined face and keystroke dynamics with information fusion at score level by adopting the action design research method. It is expected to represent a novel approach to realise the potential of multibiometric authentication in the practical scenario. The study will provide evidence of the possibility of using unsupervised learning techniques for keystroke biometrics and identify the benefits and challenges which contribute an essential element to apply multibiometric authentication.

2 Related works

Authentication is the important component of digital identity's security; therefore, authenticating a student, before and during the exam is a critical element of online exams. The authentication process usually takes place before conducting the exam (Shdaifat et al. 2020), creating an opportunity for the examinee to change after logging into the online exam or someone using another's username and password to complete the exam. Preventing cheating during the exam is also an important factor in maintaining academic integrity. Compared to traditional paper-based exams, cheating during online exams has rapidly increased due to the absence of the presence of a proctor (Shdaifat et al. 2020). It is possible that the student could use prohibited devices to help complete the exam, such as using a mobile device.

Authentication and verification processes can be categorised into three types: knowledge-based, possession-based, and biometric-based authentication. The traditional authentication methods are based on either knowledge or possession. Knowledge-based authentication uses the information known and remembered by a person such as a username and a password (Srivastava and Sudhish 2016). It is the most widely used approach to authenticate users, but it provides low-levels of security due to the high dependency on the knowledge that is susceptible to collusion and impersonation (Okada et al. 2019). A possession-based authentication approach uses an object that the user physically possesses, but this is an unpopular online authentication method because it can be stolen or copied by other examinees (Asep and Bandung 2019). Due to the challenges that knowledge and possession-based authentication approaches face, biometric-based authentication, which uses physiological and behavioural characteristics of a user are becoming increasingly used.

There exist many attempts to use biometric authentication for online exams; however, the authentication of users is only required statically at login time. One-time authentication allows access to the online exam system for an entire session; therefore, it cannot prevent cheating or impersonation from occurring during the exam (Ghizlane et al. 2019). Continuous authentication using biometric technologies has become an emerging area of research to ensure the integrity of online exams as users must be continuously verified during the exam. Continuous authentication using a number of biometric tactics has suggested facial recognition, keystroke dynamics, iris or mouse movement (Annamalai et al. 2018; Sheela and Vijaya 2010; Stylios et al. 2016). For example, continuous keystroke biometric authentication was proposed in online exams under restricted settings which do not allow users to press the backspace or delete the word typed (Flior and Kowalski 2010). In a real-world deployment, it is not possible to take an online exam without typos; hence, the restricted environment is a significant limitation. Additionally, there was no evaluation process for Flior and Kowalski's (2010) proposed system.

Apart from using one biometric trait for the continuous authentication process, a number of multimodal biometric approaches which combine more than one biometric feature are available. Depending on what processes or inputs are used for the fusion of methods and the levels of confidence score, there are different ways of combining the biometrics to perform the authentication.

There are studies on the fusion of face recognition and keystroke dynamics by combining the results generated from each modal (Gupta et al. 2015; Srivastava and Sudhish 2016). Gupta et al. (2015) proposed the combination of using face recognition and keystroke dynamics, but there is a lack of discussion on its feasibility of the continuous user verification under online exam condition and the information fusion for the decision. Another study (Srivastava and Sudhish 2016) combined the results generated from each modal (face and keystroke dynamics) at the score level to have a single score. Still, it used a virtual identity with a small number of subjects to evaluate the system with a lack of exploration of its feasibility. There is a study proposed combining three biometrics: keystroke dynamics, mouse movement and facial recognition to not to rely on a closed-world assumption for identity verification (Traore et al. 2017). In this study, the outputs were presented through separate authentication events, not by fusing the scores of the different modalities to make an overall decision.

Even though there are studies which suggest the combination of face and keystroke dynamics for continuous user verification, there is a lack of discussion of its effectiveness under the real-word deployment. The study will explore its effectiveness by implementing the system under online exam environment in University. As it was pointed by Gupta et al. (2015), the study will expand the keystroke verification by using the support vector machine (SVM) algorithm which is unsupervised learning techniques to make the module more robust. Additionally, the study explores the score fusion methods by using the decision tree combined weighted sum classification approach.

2.1 Proposed approach

The research will follow the action research design to generate prescriptive design knowledge through the building and evaluating ensemble IT artifacts in an organisational setting. This research proposes a multibiometric system, which uses facial recognition and keystroke dynamics for continuous authentication in online exam platforms. It is expected to provide a high accuracy authentication system specifically for online exams as the combination of the biometric modalities give better precision than any single biometric authentication model (Giot et al. 2010). The fusion will be completed at the score level; hence, the result of facial image and keystroke dynamics are compared with the corresponding template, then the outcome of the comparison will be combined to obtain a single fused score for overall decision making. The performance of the proposed system will be evaluated under two different categories: 1) accuracy and 2) operating cost. The accuracy will be examined based on False Accept Rate (FAR), False Reject Rate (FRR) and Enrolee False Accept Rate (EFAR) and the operating cost will be evaluated based on the execution time and CPU/memory usage from the start and the end of the authentication.

The proposed model assumes that all students from different locations are connected with the online exam platform by mobile, tablet or computer devices. The online exam authentication system will consist of three phases: 1) Registration, 2) Login, and 3) Online exam (Figure 1).

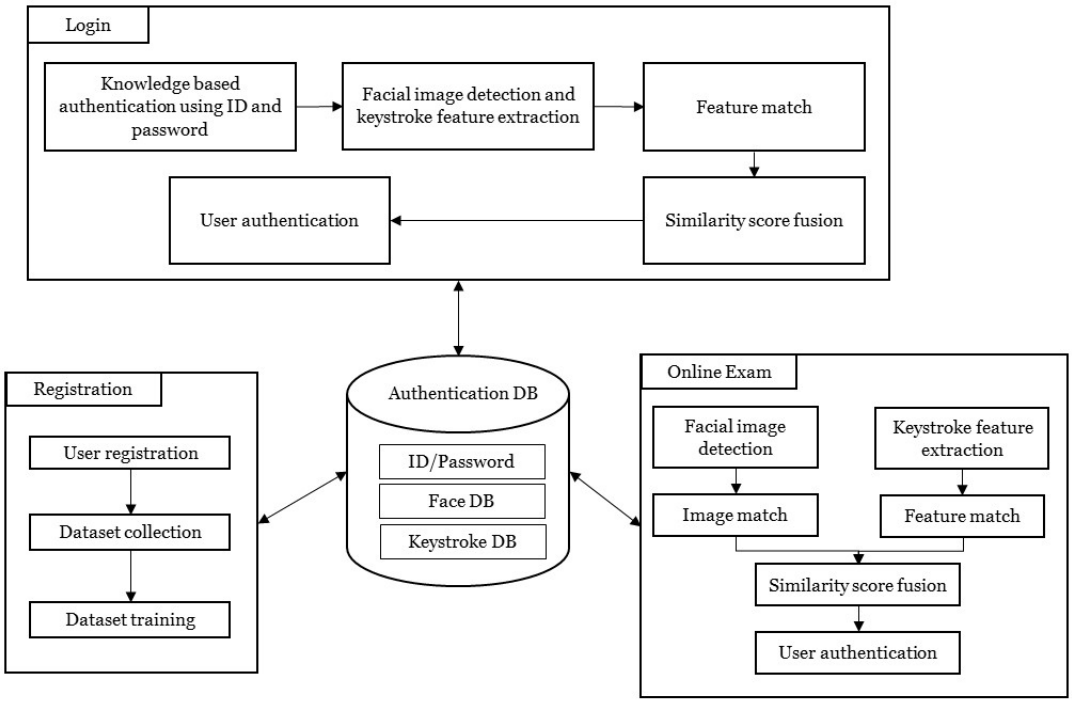

Figure 1: Continuous user authentication system for online exams

The user verification process will be executed immediately when a user logs into the exam session. Continuous authentication will be ceased when the user ends the exam. When cheating during the exam is detected due to meeting certain conditions, the exam will be reported to the supervisor by generating an alert message for further investigation. All detected cheating behaviour would be documented in the system as evidence. When one of the following condition is satisfied, it will be considered cheating: 1) the face detected and keystroke dynamics during the exam do not satisfy the target threshold, 2) multiple individual faces detected and 3) no face is detected during the exam.

At registration, a student will need to register their identity to enter into the system with the user face capture, and collection of the user's keystrokes. The user's face will be captured by the system, inviting them to look up, left and right. After the image captured, the user will be required to type a number of short sentences to collect user dynamic keystrokes.

The process of logging into the online exam platform will require two steps. The first is a traditional authentication using ID and password. When a user types their ID and password, the system checks whether the provided ID and password match with the one used to register, and it checks the keystroke dynamics with the collected typing behaviour. Once it is verified, the second step directs the user to provide the facial image through the camera. The captured image is then processed to verify the user. Through this process, the user's face image and keystroke dynamics will be updated and used as a dataset for the incremental training process, which will be executed in the server on a regular basis.

During the exam session, facial image and typing behaviour will be repeatedly captured at set times to ensure the continuous authentication of the user. This process can be divided into three stages: 1) face authentication, 2) keystroke authentication and 3) decision making (Figure 2).

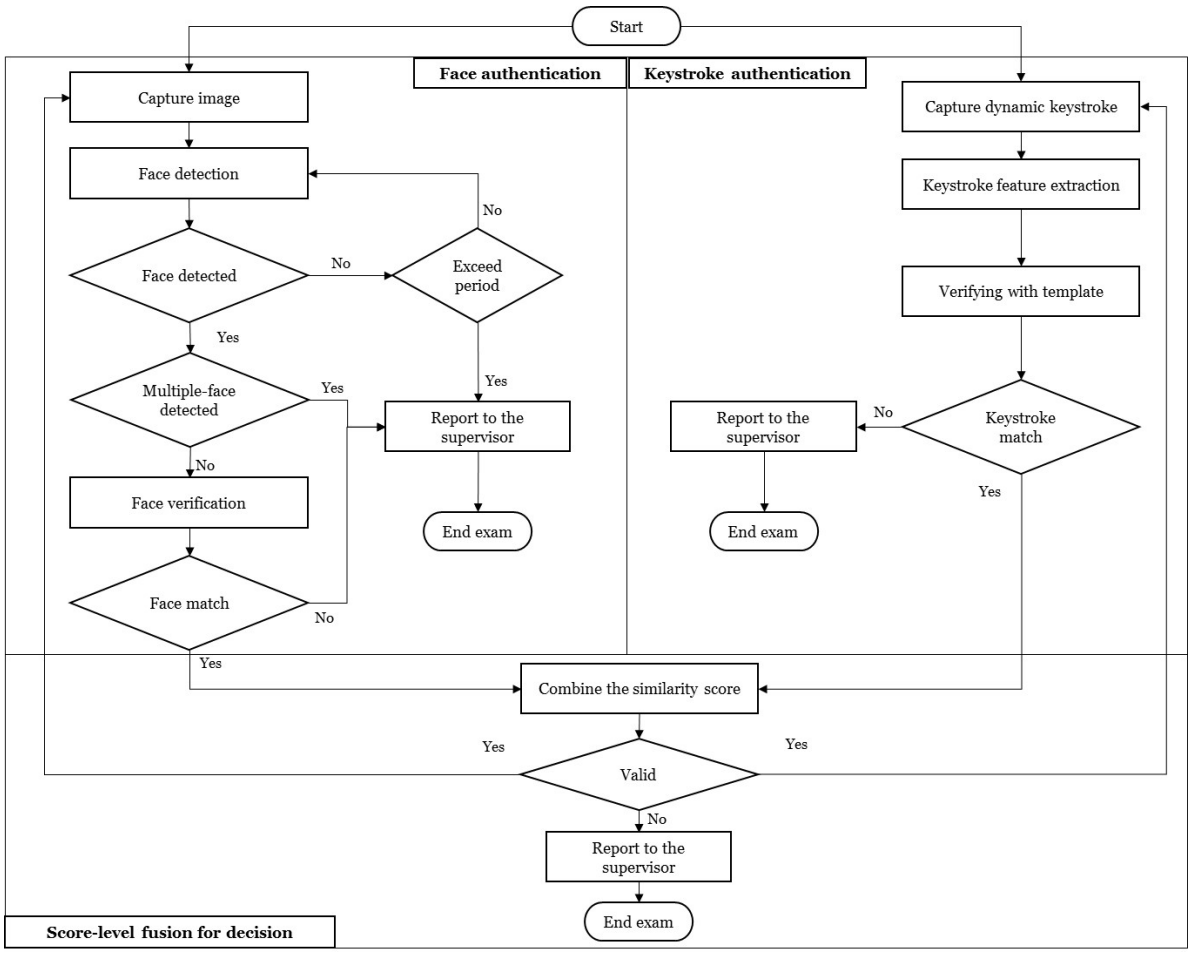

Figure 2: Flowchart of proposed continuous user authentication process with multibiometric factors (face and keystrokes) during exams

2.2 Unimodal systems

The continuous face detection and verification process will be undertaken by using the EigenFace approach because of its speed, learning capacity and simplicity (Srivastava and Sudhish 2016). The EigenFace approach uses Principal Components Analysis (PCA) to calculate a set of eigenvectors from several parts of a collection of training face images in order to reduce the dimensionality of the data by acquiring as much variation as possible in the original dataset. If the face is not detected, the period of no-face will be measured to determine whether it continues or drops the continuous authentication process. When the system detects multiple faces, then it is conceived as potential cheating behaviour. If a single face is detected, the detected face will be verified based on the EigenFace approach.

For the keystroke dynamics, the SVM approach will be used as a classifier engine for the authentication due to its high recognition rate and efficient processing (Çeker and Upadhyaya 2016). During the registration and login phases, the keystroke data will be collected to calculate the duration between keystrokes (known as flight times) and the duration that a key pressed (known as dwell times) accordingly. A one-class SVM will be trained for each user separately using only the digraph information that belongs to the individual users as one-class SVM does not require negative-data points. All users will be cross-checked with each other to discover how one-class SVM is successfully identified hyperplane that separates a user from another.

2.3 Information fusion at the score level

The score obtained from the two different modalities may not belong to the same distribution or range to combine directly; therefore, score normalisation is essential for transforming these scores from different systems into a common domain before combining them (Srivastava and Sudhish 2016). Min-max normalisation will be applied to normalise the score with the range [0,1]. The minimum and maximum scores are either known or are classified from the given data. The scores will then be fused based on classification and combination, as was suggested by (Aizi and Ouslim 2019). It applies a

cutting method of the score range into zones of interest. It cuts the score range from 0 to 1 into three zones of interest. Zone 1 is a certainty zone where the user is identified. Zone 2 is the uncertainty zone where the identification is not sufficiently reliable, and zone 3 is the certainty zone where the identification of the user failed. In order to cut these zones, the test will be performed on both modalities by plotting the False Accept Rate (FAR), False Reject Rate (FRR) and Enrolee False Accept Rate (EFAR) according to different decision thresholds from 0 to 1. The clustering method, such as K-means will be applied to cut the score range from 0 to 1 into three zones of interest for both modal systems. Once the distribution of identification scores of keystrokes and face recognition systems is studied through the training phase, a fusion approach will be applied to decide whether the user is identified or not based on the zones of interest. This approach will follow the classification by decision tree together with the weighted sum.

3 Conclusion

This paper provides an overview of preliminary research into a multibiometric continuous authentication system using two modalities, keystroke dynamics and face recognition system in online exams. Two types of authentication are proposed to authenticate students for the online exam platform. The system requires a traditional login method using username and password, in addition to the biometric authentication using face recognition and keystroke dynamics to verify the user's identity. Both face recognition and keystroke dynamics will be verified continuously during the exam to prevent impersonation. Among the various existing level of fusion, we chose fusion at the score level as the scores coming from different modalities are rich in information.

The concepts and research design presented in this paper will provide novel evidence for realising the potential of multibiometric authentication by offering an examination of its effectiveness in the practical scenarios outlined here. It will provide supportive evidence whether the suggested approach can be applied in the different biometric systems to improve the performance as it was demonstrated in the previous study. The benefits and issues address throughout the experiment will provide insight into the application of multibiometric authentication under any practical environment. As a continuation of this work, we will develop a prototype to test the suggested model in this article and perform a test on real multimodal databases that contain a high number of individuals and biometric samples. However, the study contains the limitation as we limit our experimental focus on the online exam using a computer. Therefore, the results of the study could not be applicable to the mobile platform. As such, this aspect paves for future research, intending to lighten the multibiometric authentication, which can be suitable for mobile.

4 References

Aizi, K., and Ouslim, M. 2019. "Score Level Fusion in Multi-Biometric Identification Based on Zones of Interest," *Journal of King Saud University - Computer and Information Sciences*).

Annamalai, P., Raju, K., and Ranganayakulu, D. 2018. "Soft Biometrics Traits for Continuous Authentication in Online Exam Using Ica Based Facial Recognition," *I. J. Network Security* (20), pp. 423-432.

Asep, H. S. G., and Bandung, Y. 2019. "A Design of Continuous User Verification for Online Exam Proctoring on M-Learning," *2019 International Conference on Electrical Engineering and Informatics (ICEEI)*, pp. 284-289.

Çeker, H., and Upadhyaya, S. 2016. "User Authentication with Keystroke Dynamics in Long-Text Data," *2016 IEEE 8th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, pp. 1-6.

Dasgupta, D., Roy, A., and Nag, A. 2016. "Toward the Design of Adaptive Selection Strategies for Multi-Factor Authentication," *Computers & Security* (63), pp. 85-116.

Flior, E., and Kowalski, K. 2010. "Continuous Biometric User Authentication in Online Examinations," *2010 Seventh International Conference on Information Technology: New Generations*, pp. 488-492.

Ghizlane, M., Hicham, B., and Reda, F. H. 2019. "A New Model of Automatic and Continuous Online Exam Monitoring," *2019 International Conference on Systems of Collaboration Big Data, Internet of Things & Security (SysCoBIoTS)*, pp. 1-5.

Giot, R., Hemery, B., and Rosenberger, C. 2010. "Low Cost and Usable Multimodal Biometric System Based on Keystroke Dynamics and 2d Face Recognition," *2010 20th International Conference on Pattern Recognition*, pp. 1128-1131.

Gupta, A., Khanna, A., Jagetia, A., Sharma, D., Alekh, S., and Choudhary, V. 2015. "Combining Keystroke Dynamics and Face Recognition for User Verification," *2015 IEEE 18th International Conference on Computational Science and Engineering*, pp. 294-299.

Okada, A., Whitelock, D., Holmes, W., and Edwards, C. 2019. "E-Authentication for Online Assessment: A Mixed-Method Study," *BJET* (50), pp. 861-875.

Shdaifat, A. M., Obeidallah, R. A., Ghazal, G., Abu Sarhan, A., and Abu Spetan, N. R. 2020. "A Proposed Iris Recognition Model for Authentication in Mobile Exams," *International Journal of Emerging Technologies in Learning (iJET); Vol 15, No 12 (2020)*).

Sheela, S. V., and Vijaya, P. A. 2010. "Iris Recognition Methods - Survey," *International Journal of Computer Applications* (3).

Srivastava, S., and Sudhish, P. S. 2016. "Continuous Multibiometric User Authentication Fusion of Face Recognition and Keystoke Dynamics," *2016 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)*, pp. 1-7.

Stylios, I. C., Thanou, O., Androulidakis, I., and Zaitseva, E. 2016. "A Review of Continuous Authentication Using Behavioral Biometrics," in: *Proceedings of the SouthEast European Design Automation, Computer Engineering, Computer Networks and Social Media Conference*. Kastoria, Greece: Association for Computing Machinery, pp. 72–79.

Traore, I., Nakkabi, Y., Saad, S., Sayed, B., Ardigo, J., and Quinan, P. 2017. "Ensuring Online Exam Integrity through Continuous Biometric Authentication," in *Information Security Practices: Emerging Threats and Perspectives,* I. Traoré, A. Awad and I. Woungang (eds.). Switzerland: Springer, pp. 73-81.

Velásquez, I., Caro, A., and Rodríguez, A. 2018. "Authentication Schemes and Methods: A Systematic Literature Review," *Information and Software Technology* (94), pp. 30-37.

Acknowledgements

This document was adapted from the Instructions for Authors from ICIS2007 (which in turn was adapted from the AMCIS templates), PACIS 2007, ACIS 2011, ACIS 2010, ACIS 2008, ACIS2007, ACIS 2006, and the ACIS 2005 Instructions, which were an extension of the ACIS 2004 instructions, much of which was adapted from the ACIS 2003 and ACIS 2002 Instructions, which were based on the ACIS'98 Instructions (which was adopted from ACIS'97 Instructions). These in turn were adapted from an "Instructions for Authors" written by Roger Clarke. The new format, use of the Creative Commons license and support for DOIs was added by John Lamp in 2015.

Copyright