International Journal of Computer Applications (0975 – 8887)
Volume 185 – No. 32, August 2023

Validating Biometric Pattern Recognition using Tableau: An Empirical Study

S.M. Emdad Hossain
Department of Information Systems, College of EMIS, University of Nizwa, Oman

M. Raisuddin Ahmed
Militery Technical College, Muscat, Oman

Mohammad Amir
Department of Information Systems, College of EMIS, University of Nizwa, Oman

S. Arockiasamy
Department of Information Systems, College of EMIS, University of Nizwa, Oman

Sallam Fageeri
Department of Information Systems, College of EMIS, University of Nizwa, Oman

ABSTRACT
In this paper tableau is being experimented as a primary tool to validate the results of biometric pattern recognition e.g. face, ear, gait etc. which will lead to person identification. After completing the dimensionality reduction using LDA (linear discriminant analysis), we will classify every single subject using number of prominent classifiers MLP and SMO. By using tableau, we are planning to confirm the right pattern of subject ID by comparing the training dataset with the testing data. For the experiment we will use multi-dimensional data from use CASIA database. After extracting the image from the video we will reduce dimensionality of every single image before mining the Eigen vector. Finally, the Eigen vector will feed to our proposed platform as two different form of data (training and testing). Matlab is our primary language for the experiment.

Keywords
Validate, Tableau, LDA, MLP, SMO

1. INTRODUCTION
Validating an experimental result is becoming very important to get the acceptability in its users. Number innovations and inventions we find more often from research community, but not all of them coming to real existence. This might be lack of validation, lack of interest or lack of opportunity. In this research we tried to up-hold the importance the validation of achieved result in order to bring the finding to real existence. In this paper we did experiments on biometric pattern recognition to establish human identity. Even though we found significant accuracy in true-identification; still need of validation is on demand. By considering the demand, we apprentice TABLEAU as tool to validate our attained results. World prominent biometric gait dataset CASIA [4] is applied for our experiments along with PCA-LDA and two well-known classifiers (MLP-SMO). We tried to dig-down how we can validate our accomplished results in order to attain the confidence from user level stakeholder. Extensive experiments we conducted to achieve virtuous result in context of robustness. Each set of results again verified and confirmed with TABLEAU where we found enough confidence on TABLEAU to apply in the area of pattern recognition and computer vison. We are also certain that it can be a promising tools to validate the result extracted from algorithms and classifiers. Similarly, in this research the obtained results are undoubtedly promising which will accelerate the next stage of research in the same domain.

2. BACKGROUND
Recently visual analytical platform getting more attraction to the biometric research community to visualize and to validate their achieved results in various viewpoints. Perhaps it is becoming actual needs due to number valid reasons e.g. complexity in practical implementation or not providing projected results and so on. Since then validation is taking place in these regards. In general research in biometric identity verification is a continuous task for the scholar in the same filed. Other than the general identification technique,

The performance of two (2) dimensional (2D) face recognition system depends a number of variables especially it depends on how smoothly or how accurately we can deal and mitigate problem with facial expression, various make-up and aging. Likewise, it also depends on various external factor with camera and viewpoints. Illumination difference and scene geometry are two important factors to be considered as well [6]. However, the variation in illumination and picture positions or pose are vulnerable in recognition using 2D recognition system. In these circumstances, better recognition or result can be obtained using 3D face. Because it overcomes the variation in illumination and picture positions or pose. At the same, measuring or taking care of all possible points in the facial surface was proposed again to work with 3D face recognition system [6], even though the advancement towards to real time implementation is not very much promising. It is in fact, due to cost of required tools along with its infrastructure and scarcity in workable dataset or database [7].

Consequently, the biometric security community delved into investigating alternative biometric characteristics as well as combinations of various biometric traits. Subsequently, as the global demand for automated security and surveillance products grew, there was a surge in research efforts focusing on identity verification using diverse biometric modalities [8], [9], and [10]. The importance of this research is already proven as a good number of research and development already been through. The bottom-line findings of all those studies are multiple modalities instead of single are more effective and efficient to robust identification [9], [10] and [11]. Nevertheless, the majority of these systems have primarily undergone testing in controlled laboratory environments, posing a significant challenge to replicate comparable levels of accuracy and reliability in real-world public surveillance applications. Moreover, the current array of identity authentication systems predominantly relies on modalities such as fingerprints, palm prints, faces, irises, and ear biometric traits [12], [13], and [14]. However, these modalities possess limitations when it comes to their practical applicability in public surveillance scenarios or conducting authentication from a distance.

Recently, there has been a growing demand for long-range surveillance in public areas and establishments, driven by the need to safeguard against terrorist attacks and protect public assets. Computerized video surveillance systems have emerged as the initial line of defense in various operational scenarios and applications. These systems play a crucial role in protecting assets and individuals in diverse settings, such as granting access control to civilian public spaces, facilitating financial and transaction-related activities, and ensuring stringent security measures at immigration and border control checkpoints. This approach of security and identification is opening new door when we all are talking about digital economy. Perhaps it will undoubtable empower of confidence, veracity and safety. [15], [16] and [17]. Though, Surveillance systems specifically designed for high-security environments face significant challenges when deployed in everyday civilian settings, as they struggle to cope with the uncontrolled presence of noise and non-ideal operating conditions prevalent in such environments. Subsequently the correct establishing individual credentials in question mark in some instances.

However, it is time think on the validation of the result for flat implantation to secure the benefit of specific research. By considering the drawback of existing techniques, and considering the potentiality of our proposed validation-approach we are proposing to test (tableau) We propose a novel human identification scheme that utilizes long-range gait profiles extracted from surveillance videos is proposed in our actual experiment, we aim to analyze the characteristics of multi-view gait images captured by multiple cameras. We will assess the effectiveness of infrared and visible range imagery in accurately determining a person's identity. Additionally, we will investigate the impact of multimodal fusion, efficient subspace feature extraction techniques, and classifier methods on the performance of the identification system. Furthermore, we will explore the potential of incorporating soft biometric factors, such as walking style, to enhance the accuracy and robustness of the identification systems. Finally, we will validate the result with our proposed validation tools TABLEAU. The total methodology is going to discus in the next section.

3. METHODOLOGY

Our experiments encompassed a diverse range of approaches and methods tailored to address specific objectives. To unravel patterns within the data and highlight their similarities and differences, we employed Principle Component Analysis (PCA) and Linear Discriminant Analysis (LDA) alongside MLP and SMO classifiers. PCA emerged as a potent analytical tool, particularly when grappling with high-dimensional data that lacks the luxury of visual representation. Its remarkable capability to identify data patterns enables us to compress the information by reducing the dimensions without substantial loss of vital details. Notably, PCA finds extensive utility in image compression techniques [18], functioning by organizing the rows of pixels into a one-dimensional image matrix represented by

$$X=(x1, x2, x3...N^2) \quad (1)$$

In our approach, we arranged the rows of pixels in each image sequentially to construct a one-dimensional representation. Each image possessed dimensions of N pixels in height and N pixels in width. From each image, an image vector was created. By combining all the images, we constructed a comprehensive image matrix, represented as

$$\text{Matrix} = (v1, v2, v3...vN) \quad (2)$$

This matrix encapsulated the collective information from all the images, facilitating further analysis and exploration of the data.

On the other hand, Linear Discriminant Analysis (LDA) shares a close connection with Principal Component Analysis (PCA) and factor analysis, as they all aim to discover linear combinations of variables that effectively explain the data. However, LDA stands out by explicitly focusing on modeling the differences between various classes of data. In contrast, PCA does not take class differences into account, while factor analysis constructs feature combinations based on dissimilarities rather than similarities. It's important to note that discriminant analysis and factor analysis also differ in terms of their analytical approaches. Discriminant analysis requires distinguishing between independent variables and dependent variables (also known as criterion variables), making it a non-interdependence technique. LDA proves to be effective when dealing with continuous measurements of independent variables for each observation. When categorical independent variables are involved, an equivalent technique called discriminant correspondence analysis is used [1]. In our experiment, LDA emerged as a prominent choice compared to PCA, demonstrating its effectiveness in addressing our research objectives.

Furthermore, the Multi-Layer Perceptron (MLP) represents a type of neural network architecture characterized by its feedforward structure. This means that information flows unidirectional, progressing from the input layer to the output layer in a forward manner. The MLP consists of one or more hidden layers, strategically positioned between the input and output layers. When it comes to training, the MLP relies on the backpropagation learning algorithm, which enables iterative adjustments of the network's weights to optimize performance. Renowned for its versatility, MLPs find widespread use in diverse areas such as pattern classification, recognition, prediction, and approximation. Notably, a notable strength of the Multi-Layer Perceptron lies in its capacity to effectively tackle problems that are not easily separable by linear boundaries, demonstrating its flexibility and suitability for complex tasks [2].

In addition, the Sequential Minimal Optimization (SMO) algorithm stands as an SVM classifier that leverages a distinctive learning approach. SMO effectively breaks down the overarching Quadratic Programming (QP) problem into smaller QP sub-problems, employing Osuna's theorem to ensure convergence [3]. Unlike alternative methods, SMO adopts a strategy of addressing the most compact optimization problem at each step. One significant advantage of SMO lies in its ability to analytically solve for multi-instance multipliers, obviating the need for additional matrix storage. The SMO algorithm encompasses two key components: an analytic method for determining the two Lagrange multipliers and a heuristic approach for selecting the most suitable multipliers to optimize [3].

In the context of the SVM, the relationship between y1 and y2 can be expressed as:

$$y1 \neq y2 \Rightarrow \alpha1 - \alpha2 = k \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3)$$

$$y1 = y2 \Rightarrow \alpha1 + \alpha2 = k \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4)$$

Nevertheless, it is imperative for the multi-instance multipliers to satisfy all the constraints inherent to the complete problem. The presence of a linear equality constraint dictates their placement along a diagonal line. Consequently, each iteration of the SMO algorithm endeavors to identify the optimal solution for the objective function along a specific segment of this diagonal line [3]. In the subsequent section, we delve into a comprehensive account of the experiments conducted, utilizing the powerful data visualization tool, TABLEAU, to provide intricate details and insights.

4. EXPERIMENTS AND ANALYSIS

For our experiment we used data from CASIA database. To validate all our experiments, we have gone through extensive method of validation using tableau. However, in this instance a number of experiments has been conducted to find-out the effectiveness of gait as a long-range biometric trait. The gait information captured in various viewpoints e.g. 45-degree 90-degree etc. using both infrared and thermal camera. Similarly, soft secondary biometric also combined with gait biometric trait. Father in the gait biometric trait both fast and slow walking are considered separately. Three different set of data from three different viewpoints has been tested in our first set of experiment. Visible range camera is in used for capturing all those data. Even though the data captured in a control environment, but various dimensions and strategy has applied to make it in a natural or usual behavior of walking. Table-1 and Figure1 shows the obtained result LDA-MLP, LDA-SMO, PCA-MLP, and PCA-SMO. As demonstrated in the table; experiments were mixed of individual view point and fusion of multiple viewpoints.

Table 1: Result of 3-D fusion

No	Method	Viewpoint	Recognition Accuracy
1	LDA-MLP	36 Degree	98%
2	LDA-MLP	90 Degree	84.5%
3	LDA-MLP	126 Degree	88.88%
4	LDA-MLP	3D View-Fusion	99%
5	LDA-SMO	Fusion	99%
6	PCA-MLP	Fusion	82%
7	PCA-SMO	Fusion	68%

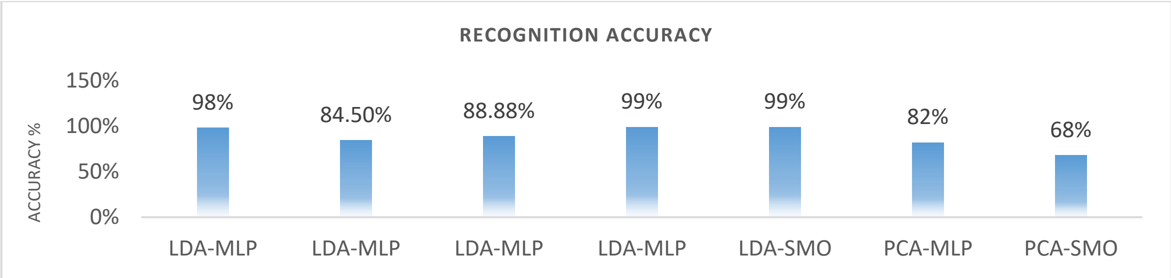

Fig1. Result of 3-D fusion

By taking close consideration of our initial results above, we conducted 2nd set of experiments to validate the accomplished results using TABLEAU. Tables below shows the results of each set of experiments.

Table 2: Validation result for LDA-SMO

Number of Clusters:	3
Number of Points:	10
Between-group Sum of Squares:	7.6602
Within-group Sum of Squares:	2.2557
Total Sum of Squares:	9.9159

In this experiment, the k chosen is 3. The total sum of square is 9.9159 and the distance between groups is 7.66. Cluster2 have the highest Eigen vectors of 5 vectors. The p value for all S groups is ~ 0.1 which is more than 0.05 (significance level). The lesser the p-value, then more the expected values of the elements of the corresponding variable differ among clusters.

5. CONCLUSIONS AND FUTURE WORKS

After successful completion of our multi-stage experiments, we came-up with a solid background to go further on TABLEAU to validate biometric pattern recognition results. In this set of experiments, we initially obtained the result of recognition using PCA-LDA as algorithms and MLP-SMO as classifiers. Both classifiers we applied with stated algorithms. After that we had 2nd set of experiment to validate the results using TABLEAU. One of significant findings to conclude with; the result we achieved using our initial combinations (algorithm-classifier), are very much similar after validating with the TABLEAU. This is a clear indication that the TABLEAU can be a powerful tool to validate biometric pattern recognition. However, further experiment and validation will be required in order to implement our claim in real-time environment, which is in fact, our future domain of this research.

6. REFERENCES

[1] Yu, H.; Yang, J. (2001). "A direct LDA algorithm for high-dimensional data — with application to face recognition". Pattern Recognition. 34 (10): 2067–2069

[2] Haykin, S. (1994). Neural networks: a comprehensive foundation. Prentice Hall PTR

[3] S.S. Keerthi, S.K. Shevade, C. Bhattacharyya, K.R.K. Murthy (2001). Improvements to Platt's SMO Algorithm for SVM Classifier Design. Neural Computation. 13(3):637-649.

[4] Get more data insights, Tableau, © 2003-2022 Tableau Software, Llc, A Salesforce Company. All Rights Reserved

[5] S. Zheng, "CASIA Gait Database," Institute of Automation, Chinese Academy of Sciences, www.sinobiometrics.com

[6] S. Berretti, A. Del Bimbo, and P. Pala, "3D Face Recognition Using Isogeodesic Stripes," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 12, pp. 2162–2177, 2010.

[7] Online Resource, “Human Action Recognition,” Advantages and disadvantages of 3D face recognition. [Online]. Available: http://www.tutorial.freehost7.com/human_face_recognition/advantages_and_disadvantages_of_3D_face_recognition.htm.

[8] A. K. Jain, “Next Generation Biometrics,” Department of Computer Science & Engineering, Machigan State University, 10-2009

[9] S. Bengio and J. Mariéthoz, “Biometric person authentication is a multiple classifier problem,” in Proceedings of the 7th international conference on Multiple classifier systems, Berlin, Heidelberg, 2007, pp. 513–522.

[10] G. Shakhnarovich and T. Darrell, “On Probabilistic Combination of Face and Gait Cues for Identification,” in Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition, Washington, DC, USA, 2002, p. 176

[11] M. N. Eshwarappa and M. V. Latte, “Bimodal Biometric Person Authentication System Using Speech and Signature Features,” International Journal of Biometrics and Bioinformatics (IJBB), vol. 4, no. 4, pp. 147–160.

[12] B. Son and Y. Lee, “Biometric authentication system using reduced joint feature vector of iris and face,” in Proceedings of the 5th international conference on Audio- and Video-Based Biometric Person Authentication, Berlin, Heidelberg, 2005, pp. 513–522.

[13] N. B. Boodoo and R. K. Subramanian, “Robust Multi biometric Recognition Using Face and Ear Images,” arXiv:0912.0955, Dec. 2009.

[14] I. Kemelmacher-Shlizerman and R. Basri, “3D Face Reconstruction from a Single Image Using a Single Reference Face Shape,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 2, pp. 394–405, 2011.

[15] G. Chetty and M. White, “Multimedia sensor fusion for retrieving identity in biometric access control systems,” ACM Trans. Multimedia Comput. Commun. Appl., vol. 6, no. 4, pp. 26:1–26:21, Nov. 2010.

[16] G. Chetty and M. Wagner, “Robust face-voice based speaker identity verification using multilevel fusion,” Image Vision Comput., vol. 26, no. 9, pp. 1249–1260, Sep. 2008.

[17] S. M. E. Hossain and G. Chetty, “Next Generation Identity Verification Based on Face-Gait Biometrics,” in IPCBEE, KualaLumpur, Malaysia, 2011, vol. 11, pp. 142–148.

[18] L. I. Smith, “A tutorial on Principal Components Analysis,” Feb. 2002.