A hybrid learning scheme towards authenticating hand-geometry using multi-modal features

Mahalakshmi Basavaiah Shivabasappa, Sheela Samudrala Venkatesiah
Department of Information Science and Engineering, Faculty of Engineering, BMS College of Engineering, Bengaluru, India

Article Info

ABSTRACT

Usage of hand geometry towards biometric-based authentication mechanism has been commercially practiced since last decade. However, there is a rising security problem being surfaced owing to the fluctuating features of hand-geometry during authentication mechanism. Review of existing research techniques exhibits the usage of singular features of hand-geometric along with sophisticated learning schemes where accuracy is accomplished at the higher cost of computational effort. Hence, the proposed study introduces a simplified analytical method which considers multi-modal features extracted from hand geometry which could further improve upon robust recognition system. For this purpose, the system considers implementing hybrid learning scheme using convolution neural network and Siamese algorithm where the former is used for feature extraction and latter is used for recognition of person on the basis of authenticated hand geometry. The main results show that proposed scheme offers 12.2% of improvement in accuracy compared to existing models exhibiting that with simpler amendment by inclusion of multi-modalities, accuracy can be significantly improve without computational burden.

Keywords:
Authentication
Convolution neural network
Hand geometry
Machine learning
Palm recognition

This is an open access article under the CC BY-SA license.

Corresponding Author:
Mahalakshmi Basavaiah Shivabasappa
Department of Information Science and Engineering, Faculty of Engineering, BMS College of Engineering
PO No. 1908, Bull Temple Road, Bengaluru-560019, Karnataka, India
Email: bsmahalakshmi.ise@bmsce.ac.in

1. INTRODUCTION

Biometrics come in many different forms, for example, voice, iris, fingerprint, and facial geometry. The majority of commercial and residential applications employ hand-based biometrics the most out of all the options [1], [2]. The creation of the biometric template is the first stage in using a hand-based user authentication system. The hand picture is captured in its raw form, and then registration is carried out to perform authentication process [3], [4]. A template is essentially a digital depiction of specific qualities that is created by building a 3D model of the back, front, and palm traits [5]. However, the authentication mechanism suffers from intrinsic problems associated with the template creation process. The input image is created by manually removing the extraneous components from the acquired original image. Because of this operation, it is computationally challenging for the authentication algorithm to determine differential changes due to hand position [6], [7]. In addition, the underlying process of template registration and validation is also subjected to inaccuracies and potential flaws. This results in a small amount of physical similarity between the geometric attributes of the hand and those of other hands, which hinders the precise feature engineering. However, in recent years many solutions and schemes have been proposed to escalate the hand-geometry oriented user identification and authentication process [8], [9]. Such existing procedures are superior to conventional methods but are also observed with some limitations and pitfalls [10], [11]. The prime problem
statement is to ascertain a potential feature extraction while analyzing hand geometry because of the variable nature of the hand geometry in perspective of different modalities. In the field of biometrics, there is a trend towards multiple approaches to user authentication systems [12], [13]. However, they have a lot of obvious limitations on applications. This factor acts as a motivation to design a novel mechanism of authenticating a hand geometry considering multi-modal perspective.

Therefore, this paper introduces a unique concept of multi-modal aspects of the hand-geometry-based user authentication system. An idea is being proposed which can be used to authenticate a person using the hand geometry features of the palmer side of his hand. Hand geometry-based authentication is proven to be less computationally intensive compared to other methods used for biometric authentication. The approach which is being proposed in this paper uses two different algorithms to achieve the results. They are i) the convolution neural network (CNN) algorithm for extraction of features from images and ii) the Siamese algorithm for recognition of the person’s identity. Hand geometry recognition is achieved by measuring several factors of the hand. Some points are noted on the hand, which are called landmarks. These landmarks provide an essential insight into the posture of the hand. The CNN algorithm used in the system will extract these landmarks and detect the pose and posture of the hand. Once the carriage is known, the rest of the hand geometry features are removed using the method proposed in [14], and the hand geometry is compared in the database. The novelty of the proposed system is that in this method, both landmarks are recognized, the pose of the hand is detected first, and then the geometry features are considered. When more metadata is available, it will become easy to identify the person. The importance of metadata along with biometric information always increases accuracy. The metadata of the hands considered here are gender, skin tone, and pose. It is also known that gender recognition and authentication yielded better results than just a recognition algorithm.

The literature is rich in the context of authentication using hand-geometry features where a variety of solutions and approaches are presented. This section carries out a brief analysis of recently published literature and explores potential research problems. A manifold biometric system is introduced by Gupta and Gupta [15] who utilized combination of multiple traits such as flat fingerprints, palm dorsal vein and hand geometry for an authentication process. These features had been attempted to fuse to boost the reliability of authentication process with higher rate of user acceptability without taking much time in acquisition. In the same line of research, Jaswal et al. [16] carried out feature-level fusion of flat fingerprints, palm dorsal vein and hand geometry. Initially, the segmented region of interest is subjected to geometrical correction and illumination enhancement and principal component analysis is adopted to determine the similarity and feature differences to perform classification. Mohammed et al. [17] suggested multi-featured based authentication system to authorize persons. A total of 21 features of the right hand were extracted using geometrical characteristics and three different types of neural network were implemented to carry out person recognition. A gender recognition specific system is introduced by Afifi [18], who had implemented supervised classifier (SVC) to extract significant features of hand image and CNN is then implemented to perform classification and recognition. Ivanova and Bureva [19] had constructed a generalized network of biometric recognition models oriented on hand geometry and vein similarity check. This work shows the idea of analyzing biometric traits through a visual sensor. Chen et al. [20] had suggested an approach of second factor authentication system against behavioral variability of user inputting. Shawkat et al. [21] had suggested an efficient recognition and verification process based on hand geometry characteristics like height, width and area of fingers. A low-pass filtering mechanism is applied to enhance hand image and 24 attributes were determined which is then fed to neural network and k-nearest neighbor (KNN) classifier. The result shows higher performance achieved by neural network compared to KNN. Oldal and Kovács [22] suggested an authentication mechanism which uses palm print and hand geometry. The feature engineering process focuses on identifying attributes related to fingertip recognition, wrist point identification and palm line extraction. This work offers a low-cost authentication system. Doroz et al. [23] had extracted biometric characteristics of hand-image and neural network based on healing mechanism is applied to perform recognition. However, this method is not cost effective as it requires extensive learning parameters and a controlled environment.

There are also other works recently carried out based on finger movement features. Alam et al. [24] had attempted to select an optimal feature to benefit classifier performance. An effective user authorization was presented by analyzing finger movements. Kumar et al. [25] introduced a security aware contactless biometric recognition system based on the features of palm vein and palm print image. The wavelet method is used to cut-down memory requirement, features are extracted using CNN and the state vector machine is then used for recognition and authentication. Ananthi et al. [26] suggested palm-vein based recognition and authentication system that utilizes fusion of curvelet multi-level and score level feature technique. Mountaines and Harjoko [27] incorporated hand geometry features with voice features for effective and robust authentication. Bartuzi and Trokielewicz [28] suggested proof-of-concept and reliable biometric authentication mechanism based on the multispectral hand features of near-infrared images and thermal.
images. Zhang et al. [29] had used a deep CNN to identify the palmprint in the biometric authentication system. Zhang et al. [30] studied schemes in the context where smartphones based biometric authentication system that uses identification of palmprint.

The prime research gap is existing system offers more sophisticated approach towards hand biometrics where emphasis is more on adoption of machine learning approach and less on effective extraction of variable features from hand geometry. The identified research problems associated with a hand-geometry-based authentication system are i) the multi-modal-based hand-geometry-based authentication system is preferred, however, the enrollment process for user images for different modalities and unique feature extraction is still computationally complex; ii) deployment of the machine learning approach and optimizing the process towards accuracy is not emphasized in the existing system towards user authentication; iii) the hand has a variety of characteristics such as gradient, texture, directionality, and many more, other modalities will also have unique qualities of their own, therefore, it is extremely challenging to develop a recognition procedure on such dispersed types of features; and iv) the relationship of one feature with another discrete element in the case of multi-modalities is not studied well, eventually leading to contextual discrepancies while working on a multi-modal hand-geometry-based authentication system. All the above-stated research problems are yet unaddressed in the existing system, and hence, the proposed method offers a solution to this. The following section discusses the proposed solution.

The core aim of the proposed study is to adopt multi-modalities of biometric traits to authenticate a person’s identity. The proposed research considers three modalities, i.e., pose, skin tone, and gender, to carry out the authentication process. The design of this system contains certain major operational blocks, as shown in Figure 1. The proposed study uses a deep learning approach for designing the recognition of pose, identification of gender, and Siamese identification. Some of the sample photos are taken from the database as reference images deployed for recognition purposes. It should be noted that these reference images are not used for training purposes.

The remaining section are structured in following sections: section 2 presents a brief review of the methodology employed in the proposed system. The experiments and performance analysis are presented in section 3. Finally, section 4 concludes the entire work carried out in this paper.

Figure 1. Block diagram of the system

2. METHOD

This section discusses the various research methods that have been implemented in order to accomplish the proposed study aim. It also bridges the research gap/problem towards using hand-geometry as simplified authentication system. Various modules implemented for this purpose are the next subsections discuss about gender recognition, pose recognition and authentication process using Siamese neural network.

2.1. Gender recognition

The proposed study uses CNN for gender recognition. The data is classified into male and female hands, followed by training the CNN. No preprocessing is being applied in this step. Along with the image, skin tone data is also considered in this scenario as skin tone also acts as a significant identifier of the person, as exhibited in Figure 2.
The proposed study performed statistical analysis using the Chi-square test to clarify the differences between male and female skin tones (other than those presented in Figure 2). This study aimed to determine whether differences in the skin tone of male and female skin were significant. The chi-square test ($\chi^2$) is a method of testing a hypothesis that can be performed when the data are a random subset of a normal sample, and the predictions and responses are inherently different. In the proposed study, the 2-test is used to analyze whether there is a statistically significant relationship between expected and actual values for one or more databases. Consider testing the variability of a process type on a small sample of $n$ product items. The sum of squares divided by a separate name value can be used to define the sample-defined $t$-number for this event. Let us consider dividing the detection of $n$ in a random sample into $k$ special classes whose numbers are $x_i$, $i=1, 2, k$. The null hypothesis, $P_i$, states that consciousness belongs to class $i$. Therefore, we have the expected number $m_i=nP_i$ for each $i$, as shown in (1).

$$\sum_{i=1}^{k} p_i = 1$$
$$\sum_{i=1}^{k} m_i = n \sum_{i=1}^{k} p_i = \sum_{i=1}^{k} x_i$$

(1)

According to the Pearson method is true in the case of the null hypothesis $n \to \infty$, the finite distribution of the quantity numerically expressed in (2) is the $\chi^2$ distribution.

$$X^2 = \sum_{i=1}^{k} \frac{(x_i-m_i)^2}{m_i} = \sum_{i=1}^{k} \frac{x_i^2}{m_i} - n$$

(2)

Pearson first encountered this issue when all cells with known numbers were expected to be large enough that all $x_i$ could be assumed to have a normal distribution. They concluded that, to a large extent, $\chi^2$ follows the Ablissa2 distribution with $k-1$ degrees of freedom. However, Pearson further considered a case in which the expected numbers were based on the parameters that were to be measured in the sample and suggested that, given that $i$ was the actual expected number and $m_i$ was the expected number, the difference is shown in (3).

$$X^2 - X'^2 = \sum_{i=1}^{k} \frac{x_i^2}{m_i} - \sum_{i=1}^{k} \frac{x_i'^2}{m_i'}$$

(3)

Usually, the outcome will be favorable and tiny enough to be ignored. Finally, Pearson claimed that the mistake in this approximation will not have an impact on practical decisions if we additionally distribute $X'^2$ as a $2$ distribution with $k-1$ degrees of freedom. The Chi-square test is performed:

- $H_0$ [Null hypothesis] = There is no relationship between skin color and gender
- $H_1$ [Alternate hypothesis] = There is a relation between skin color and gender

As it can be observed, the $p$-value is $3\times10^{-47}$ which is significantly smaller than 0.05. Hence, alternate hypothesis is accepted, and the skin color data is used while training the CNN. The structure of CNN is shown in Figure 3. Table 1 describes the properties of each layer of CNN. At the output SoftMax, the activation function is being used. The outputs and performance of this algorithm are discussed in the results section.
A hybrid learning scheme towards authenticating hand-geometry … (Mahalakshmi Basavaiah Shivabasappa)
Finally, the pose of the hand is recognized by measuring distances between points number 8, 12, 16, 20, and the angles of the fingers are considered for the pose. Once the palm detection is accomplished on the hand images, the proposed model identifies three-dimensional coordinates of hand knuckles within the specified region of the hand, representing an appropriate predictive coordinate. The trends of various hand poses are subjected to learning by the proposed model, which is better sustainable even in the case of self-occlusion. To obtain ground truth value, a manual annotation is carried out along with rendering a synthetic hand model with superior visual quality to offer better insights into hand geometry. This step is performed since the subjects were asked to open and close their fingers for data diversity while capturing the images.

### 2.3. Siamese network for authentication of a person

This is the essential part of the system, which compares and authenticates a person. Siamese neural network is a twin neural network with a different layer that will differentiate between two images if both belong to the same person or not. Siamese neural network is also called one-shot learning. The flow procedure of proposed mechanism is illustrated in Figure 6.

![Siamese network](image)

**Figure 6. Design of the Siamese neural network**

The Siamese neural network uses two parallel CNN networks with the same weights. During training, their weights vary in the same way. When both images are given to the network, it essentially finds the distance between the two images. The distance will be the height of the images belonging to a different person. It will be less if it belongs to the same person. The network essentially compares two images and extracts the data on the fly. Hence the network needs to be trained very little. The loss function used over here is a particular type of mean squared error (MSE) loss. The loss is calculated in batches in the case of Siamese neural networks. The loss is always calculated between similar images and different images. The formula for the same is shown in (4).
Here, \( y(x_1, x_2) \) is the distance function assumed to be 0 if both images are the same otherwise. The standard Adam optimizer is used in this algorithm. The output of the Siamese algorithm is always between 0 to 1. 0 represents the same image, and 1 illustrates a different image. However, the result will never be exactly 0 or 1. Hence the output is rounded to the nearest integer, and then we consider it.

3. RESULTS AND DISCUSSION

The design and development of the introduced algorithm is carried out in a python language where the dataset considered here is an 11K hands dataset. This dataset is created by Gupta and Gupta [15], and it contains the hands of a total of 190 people. Multiple photos of 4 various poses of hands of each person are taken as shown in Table 2. Metadata is collected and stored. This is known to be a better dataset compared to all other standard datasets. A brief comparison of this dataset is carried out in [15], which briefly compares the 11K hands dataset with other datasets. As it can be noticed, there are no different standard data sets that provide additional information and skin color. In this dataset there are a total of 11,076 hand images with a resolution size of 1,600×1,200. The images were captured for 190 people and their ages ranging between 18 and 75 years old. Everyone was asked to open and close their left and right fingers. Each hand has a uniform white back on the spine and sides of the palm and is placed at the same distance from the camera. The dataset comes with metadata for each image, including title ID, gender, age, skin color, and hand-crafted information like right or left hand, back or palm, and whether it contains any accessories, nail polish, or irregularities.

However, it should be noted that skin color is just a part of metadata and cannot be used for segregating male to female subject. Hence, its operation is restricted within the validation with existing dataset only. The proposed system carries out authentication using their metadata. The proposed database contains a large number of hand-drawn metadata with additional details. This database is mainly considered because of the high number of average images per article.

<table>
<thead>
<tr>
<th>Table 2. Layer properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unauthorized user</td>
</tr>
<tr>
<td>Authorized User</td>
</tr>
</tbody>
</table>

Figure 7 shows that the proposed method achieves higher accuracy than other systems. The reason why the proposed gender recognition system achieves a higher accuracy is the fact that the system uses CNN and skin tone data. The fact that the skin tone data is being considered here increases the chance of accuracy since, as shown previously, the skin tone has a high correlation with gender. Also, all three channels red, green, and blue (RGB) are considered for gender recognition in the proposed system. In the rest of the cases, the image is converted to greyscale before the recognition, and the result is shown in Table 3.

Table 3 highlights the accuracy of various methods, which shows that the proposed system is good enough to be used independently as a gender recognition system. A comparative analysis of the system with different other models is shown in Table 3. In the table, it can be noticed that the performance accuracy of dorsal side outnumbers the palmer side. Md. Afifi’s model [18] using SVM performs well when compared to
GoogleNet which uses CNN. On the other hand, comparative analysis is shown in Figure 8 and our proposed model shows remarkable accuracy when compared to other models.

![Accuracy chart](image)

**Figure 7. Accuracies of various methods**

<table>
<thead>
<tr>
<th>Method</th>
<th>Palm side</th>
<th>Dorsal Side</th>
</tr>
</thead>
<tbody>
<tr>
<td>GoogleNet (CNN)</td>
<td>0.89</td>
<td>0.889</td>
</tr>
<tr>
<td>Md. Afifi (CNN) [18]</td>
<td>0.874</td>
<td>0.910</td>
</tr>
<tr>
<td>Md. Afifi (SVM) [18]</td>
<td>0.942</td>
<td>0.973</td>
</tr>
<tr>
<td>Proposed</td>
<td>0.966</td>
<td>0.982</td>
</tr>
</tbody>
</table>

**Table 3. Outcome comparative analysis**

The accuracy here is calculated by (5),

\[
\text{Accuracy} = \frac{R}{S}
\]

(5)

where \( R \) is the number of successful trials and \( S \) is the number of trials. It is nothing but the ratio of several successful trials over a total number of trials. A trial is defined as a person trying to authenticate. If the person is authorized and the system says he is authorized and the system recognizes an unauthorized person as unauthorized, the trial is said to be successful. Or else, if the system wrongly recognizes the unauthorized person as authorized or vice versa, then the trial is unsuccessful. In the proposed method, two more metrics are considered.
Percentage of unauthorized logins ($Lu$) shows the percentage of failed trials for unauthorized people. This metric is considered since this metric measures the criticality of the system. The system should not authorize an unauthorized person. It is calculated using (6).

$$Lu = \frac{A}{B}$$  \hspace{1cm} (6)

In (6), $A$ represents the number of unsuccessful trials among unauthorized and $B$ represents total unauthorized trials. In our case, the value of this metric is 0% which means the system does not allow any unauthorized person to login.

Percentage of login failure ($Lf$) is a reverse of the above situation where the system rejects the percentage of authorized people.

$$Lf = \frac{C}{D}$$  \hspace{1cm} (7)

In (7), $C$ represents the number of unsuccessful trials among authorized, and $D$ represents the total approved trials. The value of $Lf$ turns out to be 1% which means only one percent of the time; the system rejects an authorized person.

4. CONCLUSION

Hand-based biometrics have been widely used for user authentication for more than a decade; however, there are many ways to make such a system prone to intrusive attacks. Existing studies have been reviewed to discover the evolution of various machine learning approaches toward improving the authentication mechanism of the hand-based geometry system for user authentication. However, there are multiple issues concerning the existing methods. Hence, the proposed system introduces a very form of a multi-modal biometric system that uses pose recognition, skin tone recognition, and gender recognition to authenticate the user. The proposed method uses Siamese CNN with reference photos and identity to finally authorize an unauthorized person. It is calculated using (6).

$$C$$  \hspace{1cm} (7)
Dr. Sheela S.V. Associate Professor, BMSCE., Bengaluru. She can be contacted at email: bsmahalakshmi.ise@bmsce.ac.in.

Mahalakshmi Basavaiah Shivajasappa has an experience of 21-year teaching courses in different facets of information science. She has 17 publications to her credit. She is currently guiding three research scholars in the areas of biometric authentication. She has coordinated several workshops, industry visits and extension lectures including national level championships in association with IIT Kharagpur. She has actively participated in IEEE International Advance Computing Conference-IACC 2015. She is a member of IEEE. She can be contacted at email: ssv.ise@bmsce.ac.in.

Sheela Samudrala Venkatesiah has an experience of 18 years of teaching experience to supporting her in the quest of attaining excellence in the area of technical education. Supported by a diploma in Computer Science and Engineering from the Board of Technical Education, Karnataka and a bachelor’s degree in Computer Science Engineering from Kuvempu University, Karnataka. She completed her Master of Technology degree in Computer Network Engineering from VTU in 2005. She is currently pursuing a Ph.D. program from VTU in the area of biometrics and its implementation under the able guidance of Dr. Sheela S.V. Associate Professor, BMSCE., Bengaluru. She can be contacted at email: bmsmahalakshmi.ise@bmsce.ac.in.