Transactions on Machine Learning and Artificial Intelligence - Vol. 9, No. 5
Publication Date: October, 25, 2021
DOI:10.14738/tmlai.95.10808.
Gornale, S. S., Kumar, S., & Hiremath, P. S.. (2021). Handwritten Signature Biometric Data Analysis for Personality Prediction System Using Machine Learning Techniques. Transactions on Machine Learning and Artificial Intelligence, 9(5). 01-22.

Handwritten Signature Biometric Data Analysis for Personality Prediction System Using Machine Learning Techniques

Shivanand S. Gornale
Department of Computer Science
School of Mathematics and Computing Sciences
Rani Channamma University, Belagavi, Karnataka, India

Sathish Kumar
Department of Computer Science
School of Mathematics and Computing Sciences
Rani Channamma University, Belagavi, Karnataka, India

Prakash S. Hiremath
Department of Computer Science, KLE Technological University
Hubballi, Karnataka, India

ABSTRACT

Handwritten signature has been one of the most traditionally and widely accepted behavioral personal trait in Biometric security system. It contains intrinsic features in the form of shapes and patterns that reflect various dynamic and innate behavioral traits indicating a person's soft characteristics like age, gender, Personality, handedness and many more. Person's signature or handwriting represents inherently the state of the person's mind or personality characteristics at the time of writing. In this paper, a personality prediction system is proposed, which determines different characteristics of the personality of a person based on offline handwritten signature images of that person. The experiments are carried out using own dataset which is prepared for a 13 class problem. The results of a feature level fusion and classification by random forest classifier show a promisingly significant recognition rate and validate the effectiveness against the ground truth using chi-square test at 5% significance level. The performance of the proposed method is analyzed in comparison with other similar works.

Keywords: Biometrics, Graphology, LDP, LPQ, Offline Handwritten Signature, Personality Analysis, Random Forest.

INTRODUCTION

The Biometrics is the study of automated methods of quantifying human characteristics based on permanent and distinctive physiological measurements or behavioral traits. The various biometric modalities can be broadly categorized as: (i) Physiological or Physical biometrics, (ii) Behaviological or Behavioral biometrics, and (iii) Chemical biometrics. The Physical biometrics involves measurement based on physical characteristics and anatomical characteristics of an individual and deals with the traits like fingerprints, palm prints, face, iris, retina etc..Behaviological or Behavioral biometrics is the manifestation of human activity, that

involves measurement based on the way (or action) in which certain tasks have been performed by a subject. These actions may diverge or change over a period of time which is termed as behavioral biometrics, like Body dynamics, Gait analysis, Keystroke, Voice and signature analysis, etc.. Chemical biometrics, which is still an emerging field, involves measuring chemical clues such as odor and the chemical composition of human aspects like DNA, Biological analysis, blood, saliva which are analyzed by medical experts.

Handwritten signature is one of the behavioral biometric traits which is used to authenticate the individuality of a person for a long time in the human history of documentation. It is viewed as a psych-mechanism approach achieved by the writer's hand movements according to his/her brain indications [1], and it is one of the focused research area where personality of an individual can be determined/predicted. Personality including emotions, behavior, self-esteem, anger, imaginations, honesty, fears, etc.[2]. Handwriting analysis experts perform handwriting analysis, also known as Graphology, as a part of the global analysis of characteristics such as gait, voice, speech, and many other personality traits. It enables one to derive enough information about the writer's psychology and assessment based on handwriting peculiarities. Handwriting analysis has several uses in domains such as medical diagnosis, career counselling, writer identification, staff recruiting, forensic studies, psychology, matrimonial sites, e-security, e-health and many more [3-6].

Handwriting is unique to each individual and it is called as brain writing [7]. Each neural brain pattern delivers one kind of muscular coordination which is same for each individual who has that specific personality. Handwriting discloses various aspects of an individual, such as psychological disorders, morals, hidden skills, health issues, prior experiences, and also it reveals the psycho-physiological conditions of the patient. Hence, it is used as diagnostic tool [8-9]. Graphology is the analysis of the personal characteristics of person's handwriting or signature and determines their personality characteristic or state of mind of a person and it also helps in forensic science [10]. Personality identification is the scientific methodology that understands and evaluates person's emotions, which influence the structure and pattern of his/her handwriting or signature [1]. Handwritten Signature embraces size, upper ,lower and middle zones, margins, angle/slants, speed, pressure, horizontal, vertical and even flourish strokes with multilingual texts which mostly denote first name and last name, as the description made by person, representing his/her agreement or presence on that document. The handwritten-signature contains many dynamic and inherent behavioral features, which are useful in the determination of a person's soft biometric characteristics like age, gender, ethnicity, skin color, scars, handedness, personality information and height, etc. [11-13].

Several studies have been carried out to determine/predict the personality of an individual based on various characteristics/features of a handwriting or signature including page margin, line spacing, slashes, word slope, angle, size, text density, writing speed, writing regularly [2, 14], underscore lower signature, appearance of i-dot or t-bars on the letter, curved start, ending stroke and letter spacing, etc.[3,16]. The legibility, completeness, embellishments, t and dots also play a crucial role in determining the personality of individuals [17]. In the present research work, the focus is on studying how the personality does vary based on gender and proposing a prediction model based onsupervised learning techniques.

MATERIALS AND METHODS

Related Work

With the advent of computer vision technology, personality traits identification based on offline handwritten signature images is one of the focused thrust research area in pattern recognition, Graphology and, in general, the Artificial Intelligence. Consequently, several researchers have adopted the various strategies for feature extraction and classification to predict the personality of an individual using handwritten signature images.

C. Djeddi et al. [1] have proposed work on personality traits identification through handwriting analysis. Experiments are carried out by extracting a set of writing level features and are evaluated on TxPI-u database of 418 handwritten samples that yield prediction accuracy of more than 70% for the edge and hinge feature distribution.

Hemalata et al. [2] have reviewed the works on personality detection using handwriting analysis and implemented Handwriting Analysis System using various handwriting features, such as upper zones, middle zones and lower zones and suggested some more number of features like zone baseline, size, slant, spacing, margin, pressure, f, i, speed of handwriting, which can be included to make the analysis more truthful.

Nijil Raj N et al. [3] have studied forecasting the personality based on calligraphy using CNN and MLP, in which seven features that are focused are baseline, words pace, letter size, pen-pressure, top-margin, line space, slant. These seven features are extracted from every image in the dataset, which are classified using Convolutional Neural Network (CNN) combined with Multi-Layer Perceptron (MLP). The results of CNN and MLP combined model show an accuracy of 92% is obtained.

Putu Veda Andreyana et al. [14] have performed the personality analysis based on graphology of handwriting and signature, where in two techniques, namely, German and French methods, have been implemented. The German method is carried out using the image of a person's handwriting. Firstly, the handwritten image is pre-processed and then line, word, and character / letter segmentation is done. Whereas, in French Method, it tends to analyse letter-wise and then combined one. System testing was carried out using 20 samples. From the results of the system output and expert trials, it is found that an accuracy of 80% is obtained by German method and 90% by French method.

I. Maliki et al. [15] have proposed the personality prediction system based on signatures using machine learning, where four features, namely, curve start, end streak, middle stroke, and underline, are used to determine the personality. The steps taken in developing a prediction system are model training and model testing. The method used to extract features is Principle Component Analysis (PCA) and to classify is Support Vector Machine (SVM). Based on the test results presented in confusion matrix, an accuracy value of 71% is achieved. Further, it can be concluded that machine learning can be implemented to predict personality based on signatures with good accuracy.

Vaishali R. Lokhande et al. [9] have proposed work on analysis of signature for the prediction of personality traits. Signature features like underscores below signature, appearance of dot on the letter, curved start, ending stroke and disconnected streaks were extracted. The database

of 60 signatures of 10 different persons with 6 samples of each is prepared and performance of proposed method is evaluated by splitting the signature in 5 categories left, right, upper, middle and bottom. For personality prediction, artificial neural network and structural identification algorithm are used and the accuracy rates of 100%, 95%, 94%, 96% and 92% for the five categories are obtained.

Rosso O. A. et al. [18] have proposed a new method for handwritten signature classification and verification based on descriptors stanching from time causal information theory. Shannon entropy, statistical complexity and the Fisher information have been used and evaluated over the Bandt and Pompe symbolization of the horizontal and vertical coordinates of handwritten signatures. The six features, that are easy and fast to compute, are the input to the classification technique, namely,One-Class Support Vector Machine. The results are found to be better than state-of-the-art online techniques that employ higher-dimensional feature spaces which often require specialized software and hardware.

Yasemin Bay Ayzeren et al. [22] have proposed work on emotional prediction based on both offline and online handwriting and signature, and experimented using biometric database with emotional status labels – happy, sad and stress. The database comprises samples from total 134 participants with 804 handwriting and 8040 signature samples. Random forest classifier is used for prediction, which yielded an accuracy of 55-58% for handwriting and 45-50% for signature biometrics.

Gavrilescu et al. [23] have proposed a novel approach on first non-invasive three-layer architecture based neural network to determine the five personality traits of individual using offline handwriting. The database of 128 subjects containing predefined texts and random texts is used for experimentation. A prediction accuracy 77% approx. is obtained.

Muhammad Reza Aulia et al. [21] have examined the personality identification based on handwritten signature using convolution neural networks, structural analysis with margin, dot structure, disconnected streak, and separated signature features. For symbol analysis, CNN is used for the curved start, end streak, upper streak, middle streak, shell signature and underline features. The results show that the accuracy of the structural analysis approach is 78.20-96.20% and the accuracy of symbol analysis with CNN is 98.46%. It is concluded that choosing better configuration and optimization methods can improve high accuracy.

Ghosh, Subhankar et al. [27] have proposed work on the analysis of graphology-based handwritten characters for human behavior identification. The proposed method extracts structural and geometrical features of handwritten English lowercase characters a to z. Total 5300 responses were collected from different users using an automatic system. For classification, CNN and SVM classifiers are used and achieved 86.70% accuracy.

From the literature review, it is observed that the following are the research gap/challenges due to the writing pattern complexities and the variations of writing styles.

- The personality characteristic identification using handwriting pattern/signature is an active thrust research topic but still far from the satisfaction level [1].
- Personality traits are the neurological brain pattern in human brain.Due to the neuromuscular movements,an individual has a specific personality trait [2].

- The writing style changes according to the mood of a person and it depends on the language, because the personality traits may vary for different languages [3].
- Personality is an integration of all the characteristics of individuals into a decisive unique unit, which changes constantly with the environment and it is dynamic [14].
- Signature data is abstract and each signature sample has its pattern [15].
- The context is very complex when one needs to determine the personality trait if the user writes the signature in multilingual text. Therefore, the science of signature has a little bit of uncertainty.
- It is very usual that a person's signature modifies throughout hislife time due to evolution, maturity and age of personality.
- Prediction of personality from signature is the challenging and debatable domain, due to its lack of a standard. The handwriting interpretations are being made subjectively by a trained graphologist [20].
- It is still an open challenge to catalog someone as a person who is conscientious and reckless, swift or sluggish in carrying out the activities both professionally and in daily activities [21].

From the above discussion, it is observed that most of the work is developed for specific applications and that previous works are not focused on individual's gender, multilingual signatures, etc.

The contributions in the present work are as follows:

- Exploring the handwritten signature images based on gender and age.
- Deriving the personality of a person using dictionary of graphology of 13 classes.
- An interactive prediction model for personality prediction and its validation with ground truth.

Proposed Methodology

The proposed methodology is depicted in the Figure 1.

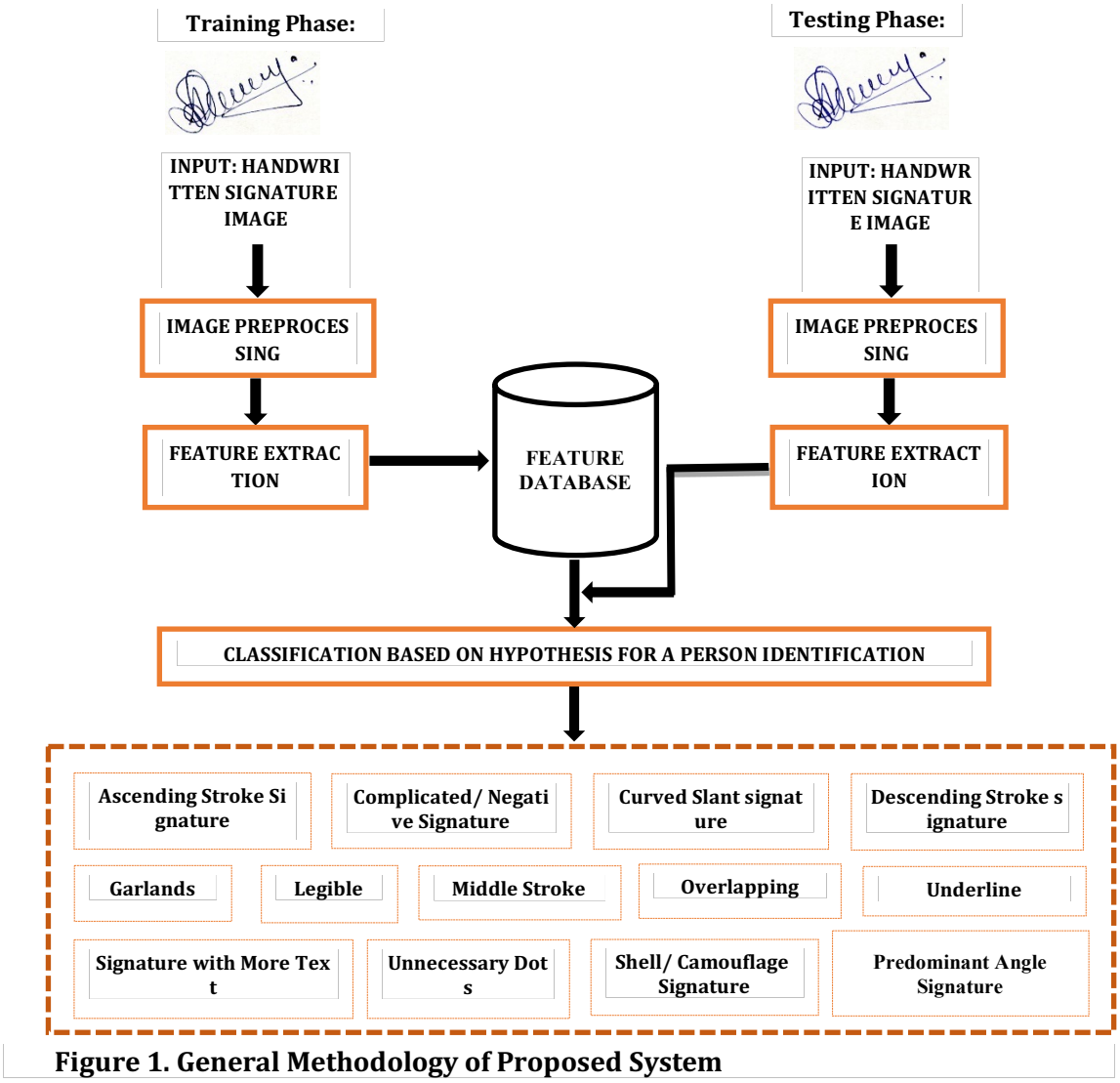


Figure 1. General Methodology of Proposed System

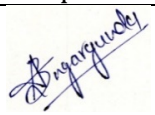
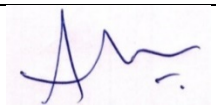
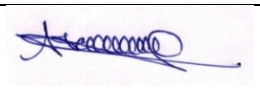
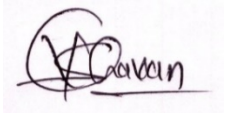
Dataset

From the above related work, it has been observed that the standard datasets, that are publicly and commercially available, do not include contributor's demographical annotations or emotion status in addition to their identification labels. With this motivation, own dataset has been created to carry out the experiments to address some of the issues raised in research gaps/challenges. The nature of this own dataset of offline handwritten signatures is described below.

- The handwritten signature samples are collected from persons of different age groups. These samples consist of multilingual scripts such as Kannada, Hindi, Marathi, and English.
- 10 signatures were recorded on a white A4 paper sheet using blue or black colored ball pen from each individuals to avoid geometrical dissimilarities in the signature.
- The recorded papers with sample signatures have been scanned using the EPSON DS1630 color scanner with a resolution of 300 DPI.
- The dataset consists of 4,790 handwritten signatures, which are acquired from 479 subjects, out of whom 250 are male and 229 female volunteers.

- Offline signature samples are classified by handwriting experts using empirical estimation technique [23] into the 13 classes and with the personality description of male and female. The 13 classes of signature images are: Predominant Angle, Ascending Stroke signature, Complicated/Negative signature, Curved Slant signature, Descending Stroke, Garlands, Legible signature, Middle Stroke signature, Overlapped signature, Shell/Camouflage Signature, Signature with more text, Signature with Under line, Signature with unnecessary Dots. Some samples with personality description are shown in Tables 1 (a) and 1(b) for male and female subjects, respectively.

Table 1(a): Signature type and description for male subjects

Signature Type	Signature Image Sample	Personality Description
Predominant Angle Signature		• It depicts aggression, • Conclusive and reluctant to change and diversity. Inability to adjust.
Ascending Stroke signature		• It is sign of ambition, optimism creativity and vitality
Complicated/ Negative Signature		• It represents distrustful personalities, • Very conflicting.
Curved Slant signature		• More Social, extroverted, affectionate and sensual with a tendency to laziness. • It indicates flexibility cordiality kindness and ability to adjustment.
Descending Stroke signature		• The person is skeptic, Pessimistic or the person is in a little struggle and feeling down
Garlands		• It represents more predisposition to extroversion, transmits spontaneity kindness and compression. • Diplomatic skills that find the exit to any situation.
Legible		• It reflects equilibrated personality with clear goals and sense of duty and responsibility.
Middle Stroke		• It represents the possessiveness of a person.
Overlapping		• It represents the invasive feelings of a person
Shell/ Camouflage Signature		• It creates self-limiting beliefs. • A person often gets defensive in public and he/she often gets defensive, over protective.

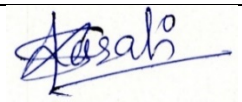
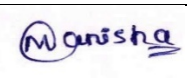
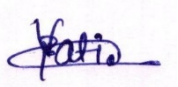
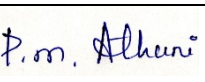
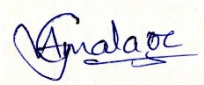
Signature With More Text		• It represents a person's dependent streak and efficiency.
UNDERLINE		• Having unique idea and thinking, need support to make decisions and have reliability in the lead
Unnecessary Dots		• It represents reflective individuals with verifying prudence, • Desire of exceeding in a perfectionist.

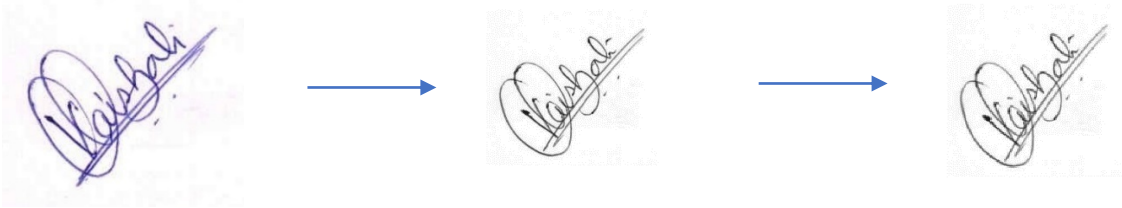
Table 1(b): Signature type and description for female subjects

Signature Type	Signature Image Sample	Description
Predominant Angle Signature		• It depicts aggression, • Conclusive and reluctant to change and diversity. Inability to adjust.
Ascending Stroke signature		• It is sign of ambition, optimism creativity and vitality
Complicated/ Negative Signature		• It represents distrustful personalities, • Very conflicting.
Curved Slant signature		• More Social, extroverted, affectionate and sensual with a tendency to laziness. • It indicates flexibility cordiality kindness and ability to adjustment.
Descending Stroke signature		• The person is skeptic, Pessimistic or the person is in a little struggle and feeling down
Garlands		• It represents more predisposition to extroversion, transmits spontaneity kindness and compression. • Diplomatic skills that find the exit to any situation.
Legible		• It reflects equilibrated personality with clear goals and sense of duty and responsibility.
Middle Stroke		• It represents the possessiveness of a person.
Overlapping		• It represents the invasive feelings of a person
Shell/ Camouflage Signature		• It creates self-limiting beliefs. • A person often gets defensive in public and he/she often gets defensive, over protective.
Signature With More Text		• It represents a person's dependent streak and efficiency.

UNDERLINE		• Having unique idea and thinking, need support to make decisions and have reliability in the lead
Unnecessary Dots		• It represents reflective individuals with verifying prudence, • Desire of exceeding in a perfectionist.

Preprocessing

The pre-processing is the first stage of the prediction system. The scanned handwritten signature images may carry spurious noise. Pre-processing is done to enhance the quality of the image. The input color image converted into gray image. Normalization is done in order to resize the signature images with the aspect ratio 150x150 that maintainsthe shape of the original signature, as shown in the Figure2.



a.Input RGB Image b. Grayscale Image c. Normalized Image

Figure 2. Preprocessing of Handwritten Signature Image

Feature Extraction

Feature extraction is the process of representing the enough shape information obtained from the Handwritten Signature image as a feature vector. Some of the feature extraction approaches are: Local Directional Patterns (LDP), Local Phase Quantization (LPQ) and, Geometric and Shape features (Area, Aspect Ratio, Maximum axis length, Minimum axis length, Eccentricity, Horizontal and vertical projection, Euler number, and Bounding Box etc.).These extracted features are then classified in order to predict the personality of person as belonging to a particular class.

Local Directional Pattern (LDP)

Firstly, the Local Direction Pattern descriptor is founded by Tasked Jabid et al. [24] in facial recognition of person. Later, it is used for signature recognition by Ferer et al. [25]. The LDP method computes,for each pixel of image, the edge response value in different directions in 3x3 neighborhood, which are used to encode the image. These methods only mark which neighbor has the analyzed characteristic on and off. To compute the LDP, we calculate the eight different directional numbers (M_0,...,M_7) of neighborhood centered at each pixel by applying convolution of Krisch compass masks with the image. The masks are shown in the Figure 3.

$$\begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix} \quad \begin{bmatrix} -3 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{bmatrix} \quad \begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix} \quad \begin{bmatrix} 5 & 5 & -3 \\ 5 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}$$

East M_0 North East M_1 North M_2 North West M_3

$$\begin{bmatrix} 5 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & -3 & -3 \end{bmatrix} \quad \begin{bmatrix} -3 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{bmatrix} \quad \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{bmatrix} \quad \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & 5 & 5 \end{bmatrix}$$

West M_4 South West M_5 South M_6 South East M_7

Figure 3. Krisch masks for computing eight directional edge responses M_0... M7.

The Local Directional Pattern code is derived by using the equation (1).

$$\boldsymbol{LDP_k = \sum_{i=0}^{7} b_i(m_i - m_k)2^i}, \boldsymbol{b_i(a)} = \begin{cases} \mathbf{1}, & \boldsymbol{a \geq 0} \\ \mathbf{0}, & \boldsymbol{a < 0} \end{cases} \quad (1)$$

Where, m_k is the k^{th} most significant directional response.

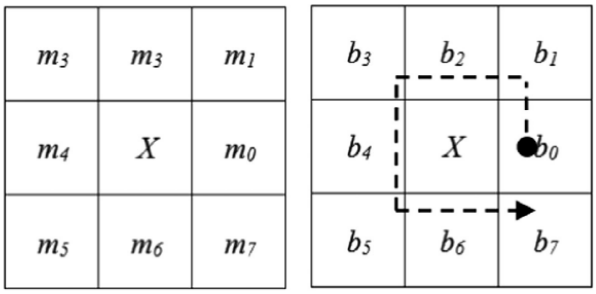
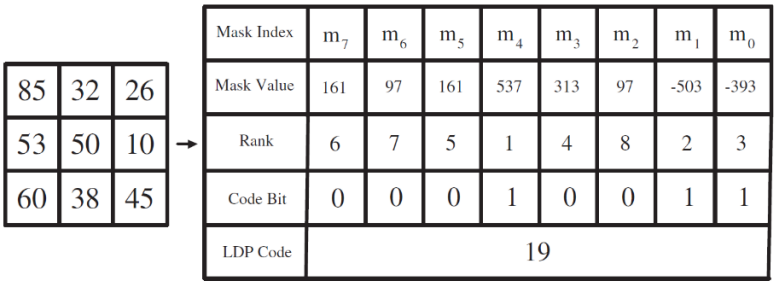


Figure 4. Represents the 8-directional edge response positions. (Left) 8-directional edge response positions and (Right) LDP binary bit positions.

By applying eight masks, we obtain eight edge response values M0, M1, M2, M3, M4, M5, M6, and M7, where each represents the edge significance in its respective direction. The values of response in all directions are not equally important. The presence of edge or corner show the high response values in particular directions. The most prominent directions are obtained in order to generate the LDP. Hence, consider the top k values |Mj| and set them to 1, and the other 8 bit LDP patterns set to 0. This allows for only 56 possible features out of the 256 [26]. The LDP code generation using the three prominent edge responses is shown in the Figure 5.



85	32	26
53	50	10
60	38	45

→

Mask Index	m_7	m_6	m_5	m_4	m_3	m_2	m_1	m_0
Mask Value	161	97	161	537	313	97	-503	-393
Rank	6	7	5	1	4	8	2	3
Code Bit	0	0	0	1	0	0	1	1
LDP Code	19							

Figure 5. Generating LDP code

The LDP produces more stable pattern in presence of noise and non-monotonic illumination changes, since gradients are more stable than grey level.

Local Phase Quantization (LPQ)

The Local Phase Quantization operation,which is proposed by Ojansivu and Heikkila [27] as a texture descriptor, is based on blur invariance property of the Fourier phase spectrum and loss of information at high frequencies. The local phase information is extracted using Discrete Fourier Transform (DFT) and more accurately, Short-term Fourier Transform (STFT). The LPQ is applied to the texture identification by computing over an M-by-M neighborhood Nx at every local pixel location x of the image f(x), and presenting the result codes as the histogram as given in equation (2).

$$F(u,x) = \sum_{y \in N_x} f(x-y) e^{-j2\pi u^T y} = W_u^T f_x \quad (2)$$

Where, Wu is the rectangular window function, and basic vector of the DFT at frequency u, and f_x is another vector containing all M^2 image samples from N_x. The scalar quantizer is defined by equation (3).

$$u(x) = \begin{cases} 1 if u_x \geq 0 \\ 0 otherwise \end{cases} \quad (3)$$

Where,the single quantizer is represented by a scalable quantizer element at a point i. u(x) is the x^{th} component of FT. Finally, the quantized coefficient values, that are presented at histogram of 256 dimensional vector, are considered as LPQ features according to the binary coding [28] given by the equation (4).

$$B = \sum_{j=1}^{L} qj^{2j-1} \quad (4)$$

This rangesbetween 0 and 255 and defines the local texture information at each image pixel positions [29].

Shape Features

The shape features are used to measure the similarities between query shapes. The features are calculated based on connected components, which transforms a contiguous and dis-contiguous data space regions into a feature space that theoretically has the same dimension as the original data space [30]. The following features are computed.

- **Area:** Area is the number of pixels in the region of an image.

$$Area = \sum_{i=1}^{n} \sum_{j=1}^{m} A[i,j] (5)$$

- **Bounding Box:**It is the smallest rectangle region, returned as 1-by-(2*Q) vector. The first Q elements are the coordinates of minimum corner of the box. The second Q elements are the size of the box along each dimension.
- **Major Axis length:**It specifies length of the major axis of an object region.

$$Majoraxislength = a + b \quad (6)$$

- **Minor Axis length:**It specifies length of minor axis of an object region.

$$Minoraxislength = \sqrt{(a+b) - f^2} \quad (7)$$

- **Euler Number:**It is the total number of objects minus the total number of holes in the objects.

$$EulerNumber = (no.ofObjects) - (no.ofHoles) \quad (8)$$

- **Eccentricity:** It is the distance ratio between the foci of the ellipse and its major axis length.

$$Eccentricity = \frac{\lambda_1}{\lambda_2}(Fociofellipse \div MajorAxisLength) \quad (9)$$

- **Aspect ratio:**It is the width height ratio of the bounding box for image object and is obtained by dividing the height by the width of the box.

$$R_i = \frac{h_i}{w_i} \quad (10)$$

Where R_i is aspect ratio of the i^{th} image object, hi is the height and w_i is the width of the i^{th} image object.

Feature Fusion

The feature level fusion is collection of the various feature patterns in one matrix, which gives a good accuracy compared to the individual feature methods [31]. To calculate complexity of larger dimensions in the feature space, the feature fusion method is utilized [28]. The offline Handwritten signatures have rich set of features, which can be combined together using fusion of features by concatenation rule to improve efficiency [32]. In the proposed work, features of Handwritten Signature are obtained from local texture descriptors, namely,LDP, Shape and LPQ, which are fused together using concatenation rule.

Classification

Classification stage is the final step of the personality prediction system.After the feature extraction, the labeled feature vectorsare used to train the model, which is then used for classifying the test feature vectors. In the experiments, different classification techniques were used, but the results were not satisfactory. To improve accuracy, we applied the ensemble classifier, namely, Random Forest Classifier that yieldedrelatively good accuracy.

Random Forest Classifier is the combination of tree structured classifiers. It can be considered as a classifier containing different classification methods or one method having several parameters [33]. Basically, it is an ensemble learning method applied for classification and regression in pattern recognition. The ensemble of classifiers $h_1(x)$, $h_2(x)$ . . . $h_K(x)$, are independent identically distributed random vectors and each tree casts a unit vote for the most popular class at input x [34].

The proposed method is represented in the following Algorithm:
Input:offline Handwritten signature image
Output: Personality Prediction, i.e. predicted class
Step 1: Input a handwritten signature image.
Step 2: Apply pre-processing of input image:grey-scale conversion and normalization of grey scale image to a size of 150x 150.
Step 3: Compute local features:Local Directional Patterns, Local Phase Quantization and Shape features individually.
Fusion of Local Directional Patterns, Local Phase Quantization and Shape features.
Step 4: Classify the fused features using Random Forest Classifier.
Step 5: Output the predicted class.
End

RESULTS

The proposed method is experimented by using on own dataset which is described in Dataset section 2.3. The dataset contains many Offline Handwritten signature images of subjects with varying gender and age. An in-house total 4790 signature images are collected from 250+229 (male and female) volunteers. Firstly, the Signature images are empirically classified and labelled by human experts in different signature type/class for both male and female signatures. The signature samples are shown in Table 1(a) and Table 1(b). Secondly, the performance of the LDP, LPQ and shape features are investigated separately. With the Local Directional Pattern (LDP), 56 features are computed and classified. Using LPQ, 256 features are extracted and classified. Next, 7 different shape features are extracted and then classified. Distribution of 60% for Training Data and 40% testing data for each class is used for classification by Random Forest classifier.

Further, feature level fusion of Local Directional Patterns, Local vector Quantization and shape features has been performed by concatenation rule of feature vectors. The fused features are stored as final vector matrix. The classification experiments were carried using single classifiers such as K-NN, SVM etc., which yielded poor results. The present problem is a 13-class problem. Hence, Random Forest classifier has been used, which yielded relatively good results. The independent performance of the LDP with highest accuracy 94.32% for male signature and 93.48% for female signature is observed. Similarly, for independent performance of the LPQ with highest accuracy 93.24% for male signature and 91.64% for female signature is observed. Also, independent performance of the Shape features with the highest accuracy 85.84% for male signature and 86.11% for female Signature is observed. Further, the fusion of different features were tested over the Signature Database. It is found that, fusion of LDP and LPQ features yielded the highest accuracy of 95.84% for male signature and 95.72% for female signature. Further, fusion of LDP and Shape Features yielded the highest accuracy of 94.40% for male and 94.36% for female signatures. Also, in case of fusion of LPQ and Shape features, the highest accuracy of 95.80% for male and 93.66% for female signatures are obtained. Finally, the fusion of LDP, LPQ and Shape features yielded enhanced accuracy of 96.64% for male signatures and 97.64% for female signatures. The Tables 2(a) and 2(b) contain the classification results of experiments and confusion matrix, from each individual features and then from fusion of features obtained by using Random Forest Classifier. The Table 3 contains the comparison of empirically estimated classification rate of Male and Female Handwritten Signature, misclassified and proposed method. Table 4 contains performance of the proposed algorithm, whichis analyzed in terms of metrics: Precision (p), F_Score (F), Recall(R) and Accuracy (A) defined by the equations(11-14), respectively.

$$\text{Precision} = \frac{TP}{TP+FP} \quad (11)$$

$$\text{F_Score} = \frac{2*Precision*Recall}{Precsion+Recall} \quad (12)$$

$$\text{Recall} = \frac{TP}{TP+FN} * 100 \quad (13)$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} * 100 \quad (14)$$

Where, TP is True Positive, TN is True Negative, FP is False Positive and FN is False Negative. The Figure 5 displays graphical representation of experimental results of feature fusion.

Table 2 (a): Confusion Matrix of Male Signature Types for Fused Features Classification

	MA	MC	MAS	MU	MCS	MG	MS	MDS	MLS	MMS	MO	MSW	MUD
MA	199	1	1	0	3	0	0	1	0	1	0	0	0
MC	1	518	0	10	2	3	3	1	0	10	4	0	1
MAS	0	0	59	0	1	0	0	0	1	1	0	0	1
MU	0	1	0	317	1	0	0	0	1	0	1	0	0
MCS	0	0	0	3	181	0	0	0	1	2	0	1	0
MG	0	0	0	1	1	127	0	1	2	1	0	0	0
MS	0	0	0	2	1	0	153	0	0	0	0	0	0
MDS	0	0	0	0	0	0	0	37	0	0	0	0	0
MLS	0	0	0	0	0	0	0	0	104	1	1	0	1
MMS	0	0	0	6	0	0	4	0	1	314	1	0	1
MO	0	0	0	1	0	0	0	0	0	0	153	0	1
MSW	0	0	0	0	0	0	0	0	0	0	0	179	0
MUD	0	0	0	0	0	0	0	0	0	0	0	0	75

where, the abbreviations in the Table 2(a) are for the 13 classes, namely, MA: Male Predominant Angle Signature, MC: Male Complicated signature, MAS: Male Ascending Stroke Signature, MU: Male Underline Signature, MCS: Male Curved Slant Signature, MG: Male Garlands Signature, MS: Male Shell/Camouflage Signature, MDS: Male Descending Stroke Signature, MLS: Male Legible Signature, MMS: Male Middle Stroke Signature, MO: Male Overlapping Signature, MSW: Male Signature with more words signature, and MUD: Male Unnecessary Dot Signature.

Table 2 (b): Confusion Matrix of Female Signature Types for Fused Features Classification

	FA	FCS	FAS	FUD	FS	FUD	FU	FSW	FDS	FG	FMS	FL	FO
FA	164	4	0	0	1	0	1	0	0	0	0	0	1
FCS	6	324	0	0	2	0	2	1	0	3	1	4	0
FAS	0	0	99	0	1	0	0	0	1	0	1	0	0
FUD	0	0	1	226	0	0	0	0	0	1	0	0	0
FS	0	1	0	0	203	0	1	0	0	0	0	0	0
FCD	0	1	0	0	0	74	5	0	0	0	0	0	0
FU	0	0	0	1	0	5	308	0	0	0	1	0	0
FSW	0	0	0	0	0	1	2	349	0	0	0	0	0
FDS	0	0	0	1	0	0	0	0	49	0	0	0	0
FG	0	0	0	0	0	0	0	0	0	85	0	0	1
FMS	0	0	0	2	3	0	1	0	0	1	177	0	0
FLS	0	0	0	0	0	0	0	0	0	0	0	126	0
FO	0	0	0	0	0	0	0	0	0	0	0	0	48

The abbreviations in the Table 2 (b) are for the 13 classes, namely, FA: Female Predominant Angle Signature, FC: Female Complicated signature, FAS: Female Ascending Stroke Signature, FU: Female Underline Signature, FCS: Female Curved Slant Signature, FG: Female Garlands Signature, FS: Female Shell/Camouflage Signature, FDS: Female Descending Stroke Signature, FLS: Female Legible Signature, FMS: Female Middle Stroke Signature, FO: Female Overlapping

Signature, FSW: Female Signature with more words signature, and FUD: Female Unnecessary Dot Signature.

In the Tables 2 (a) and 2 (b), the diagonal values correspond to correct classification and the off diagonal values correspond to misclassification. It is observed that the misclassifications are very less.

Table 3: Comparison of classification results of proposed method with manual methodfor each class

Signature Type	Manual Classification		Proposed method		Classification error		Accuracy (%)	
	Male	Female	Male	Female	Male	Female	Male	Female
Predominant Angle Signature	200	170	199	164	01	06	99	99
Ascending Stroke signature	060	100	59	99	01	01	99	99
Complicated/ Negative Signature	520	330	518	324	02	06	99	98.18
Curved Slant signature	190	080	181	74	09	06	95.26	92.5
Descending Stroke signature	040	050	37	49	03	01	92.5	99.5
Garlands	130	090	127	85	03	05	97.69	94.44
Legible	110	130	104	126	06	04	94.54	96.92
Middle Stroke	330	180	314	177	16	03	95.15	98.3
Overlapping	160	050	153	48	07	02	95.62	96
Shell/ Camouflage Signature	160	210	153	203	07	07	95.62	96.66
Signature With More Text	180	350	179	349	01	01	99	99
UNDERLINE	340	320	317	308	23	12	93.23	96.25
Unnecessary Dots	080	230	75	226	05	04	93.75	98.26

Based on the experimental results shown in the Table 3,it is observed that the classification results of the proposed method are in close agreement with the ground truth (manual results) and the personality prediction of a person is performed with good accuracy. The Table 4 displays the classification performance of individual features and the feature fusion using Random Forest Classifier. Its graphical form is displayed in the Figure 6. It is observed that the feature fusionShape+LDP+LPQ outperforms other combinations.

Table 4: The performance analysis of the individual features with that of the fused features using Random Forest Classifier

Features	Random Forest Classifier							
	Male Signature				Female Signature			
	Precision	Recall	F_Score	Accuracy	Precision	Recall	F_Score	Accuracy
LDP	0.9349	0.9548	0.9448	94.32%	0.9335	0.9379	0.9357	93.84%
LPQ	0.9237	0.9412	0.9323	93.24%	0.9102	0.9122	0.9112	91.65%
Shape	0.8417	0.8612	0.8514	85.84%	0.8457	0.8583	0.8519	86.11%
LDP+LPQ	0.9534	0.9689	0.9611	95.84%	0.9544	0.9550	0.9555	95.72%
Shape+LDP	0.9372	0.9525	0.9448	94.40%	0.9441	0.9305	0.9373	94.636%
Shape+LPQ	0.9496	0.9661	0.9578	95.80%	0.9337	0.9331	0.9325	93.66%
Shape+LDP+LPQ	0.9618	0.9727	0.9672	**96.64%**	0.9698	0.9726	0.9754	**97.46%**

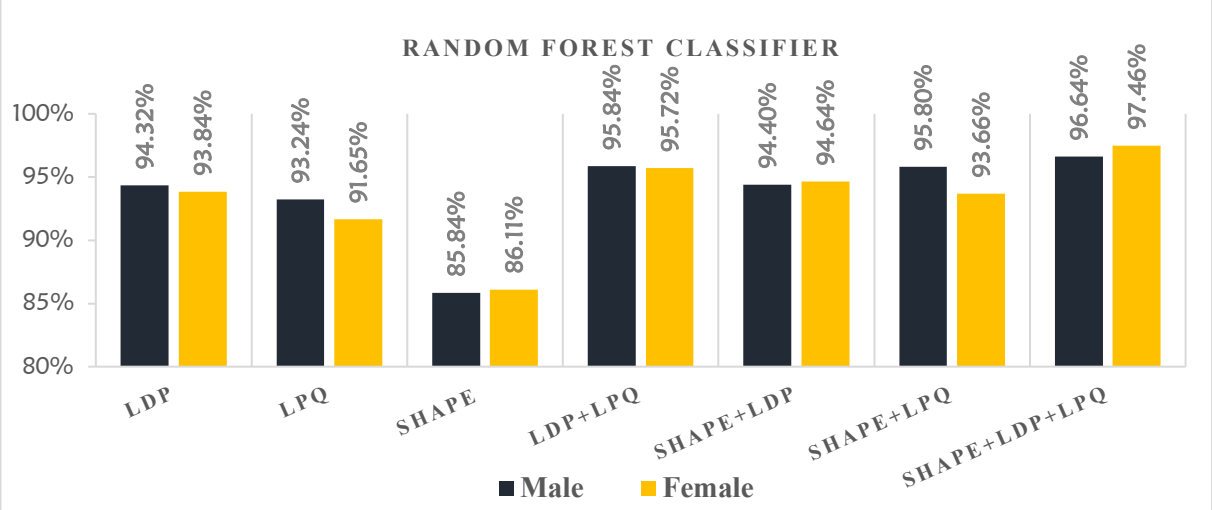


Figure 6. Classification performance of the individual features and fused features

Statistical Test of Significance

Statistical test of significance is performed to determine whether the outcomes of the experiments are statistically significant or not. It is the procedure for comparing the observed data with a claim of expected data. In the present work, Chi-Square Test is performed at 5% significance level to validate the statistical inference.

Chi-Square Test

The Chi-Square Test is used for testing the relationship between the categorical variables of the experiment. In this test, the null hypothesis, the alternative hypothesis and the degrees of freedom are stated as below:

- Null Hypothesis (H_0): There is close significant association between outcomes of the proposed algorithm and that of the manual classification.
- Alternative Hypothesis (H_a): There is no significant association between outcomes of the proposed algorithm and that of the manual classification.
- Degree of Freedom (df) =12, the critical value of $\boldsymbol{\chi^2}$with df=12 at 5% significance level is 21.03(from the chi-square table).
- If $\boldsymbol{\chi^2}$<21.0206, Accept H_0 and reject H_a, else vice versa.

The details of the Chi-Square test related to manual classification and proposed algorithm are presented in the Table 5. The value of the Chi-Square statisticis computed as given below:

$$Chi - Square(\chi^2) = \sum \frac{(f_o - f_e)^2}{f_e} = 4.11$$

The computed Chi-Square statistic value is observed to be lower than the critical value from the chi-square table. Hence, the Null Hypothesis is accepted and the alternative hypothesis rejected. It indicates that there is significant association between the proposed algorithm and the annotations provided by manual classification. The comparative analysis of manual classification and proposed method represented graphically in the Figure 7.

Table 5: The Chi-Square test on the Results of Proposed Method and Manual Classification

Signature Type	Manual Classification (Male + Female) x_i	Proposed Method (Male + female) y_i	Total (MC+PM) $x_i + y_i$	Expected values of MC f_e	Observed values of PM f_o	$\chi^2 = \frac{(f_o - f_e)^2}{f_e}$
Predominant Angle Signature	370	363	733	372	361	0.3252
Ascending Stroke signature	160	158	318	161	157	0.099
Complicated/ Negative Signature	850	842	1692	859	833	0.786
Curved Slant signature	270	255	525	266	259	0.184
Descending Stroke signature	90	86	176	89	87	0.044
Garlands	220	212	432	219	213	0.164
Legible	240	230	470	239	231	0.267
Middle Stroke	510	491	1001	508	493	0.4429
Overlapping	210	201	411	209	202	0.23445
Shell/ Camouflage Signature	370	356	726	368	358	0.271

Signature With More Text	530	528	1058	537	521	0.476
UNDERLINE	660	625	1285	652	633	0.553
Unnecessary Dots	310	301	611	310	301	0.2612
Total	$\sum x_i = 4790$	$\sum y_i = 4648$	$\sum(x_i + y_i) = 9438$			$\sum \chi^2 = 4.1137$

Where,

- x_i= manual classification values
- y_i=proposed method values
- $\sum x_i$= Total of manual classification
- $\sum y_i$= Total of proposed method
- $\sum(x_i + y_i)$= Grand total of proposed and manual method
- f_e= Expected value

$$f_e = (\sum x_i) * (x_i + y_i)/\sum(x_i + y_i) \quad (15)$$

- f_o= Observed value

$$f_e = (\sum y_i) * (x_i + y_i)/\sum(x_i + y_i) \quad (16)$$

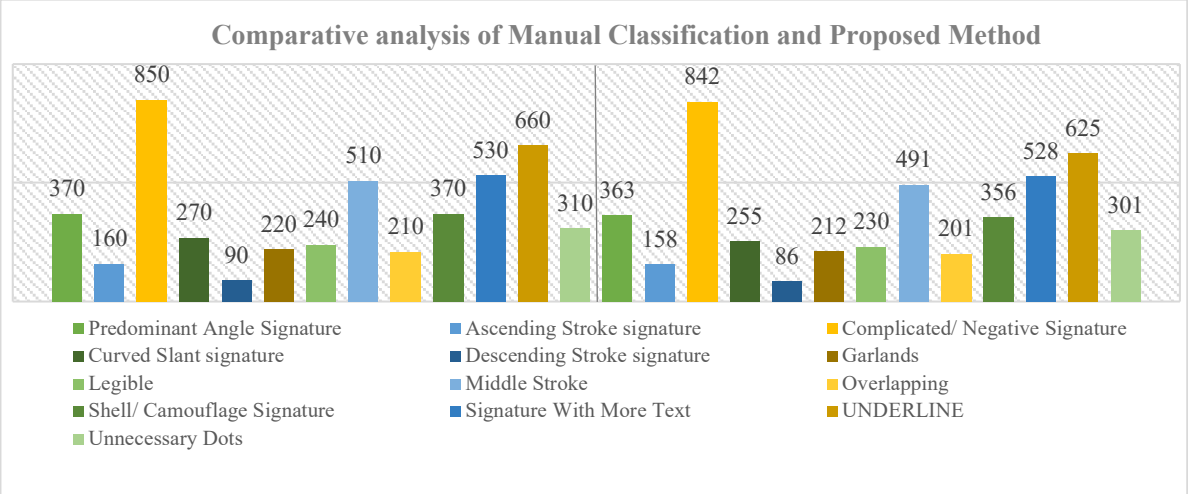


Figure 7. Graphical Representation of Results of manual Classification and Proposed Method

Comparative Analysis

The performance comparison of the proposed method with othermethods found in the related literature is summarized in the Table 6 that shows a summary of the related works in the literature, features/methodologies, classification techniques and accuracy and datasets. It is observed that the proposed method based on feature fusion and random forest classifier has yielded a comparatively higher accuracy for personality prediction than other methods.

Table 6: Comparative Analysis

Authors	Features	Database	Classifiers	Accuracy
C. Djeddi et. al. [1]	Edge-Hinge distribution Textural features	TxPI-u Database 543 handwritten samples	Artificial Neural Networks	70%
Nijil Raj N et. al [3]	Seven specific calligraphic features (Baseline, Top margin, Letter Size, Line Spacing, Word Spacing, Pen pressure, Slant)	IAM- Dataset 1 and IAM-DATASET 2	Convolutional Neural Network and Multi-Layer perceptron	92%
Putu Veda Andreyana et. al[14]	Graphology traits: margin, baseline, Size, space	20 handwritten samples	Expert System	German method 80% and French method 90%
I Maliki et. al[15]	Principle Component Analysis (PCA)	530 data signatures with 53 respondents	Support Vector Machine (SVM)	71%
Vaishali R. Lokhande et. al. [16]	Projection based feature extraction	60 signatures	Back Propagation Neural Network	94% curved start, 96% ending stroke, 100% for single ascending underline, 92% for streak disconnected and 95% for appearance of dot on signature
Yasemin Bay Ayzeren et. al. [19]	Online signatures with time dependent features	804 handwriting and 8040 signature biometric samples.	Random Forest Classifier	55-58% for handwriting and 45-50% for signatures
Gavrilescu et al. [20]	Handwriting features: Baseline, lowercase letter, connecting strokes, word slant	128 subjects containing predefined texts and random texts	Neural Network	77%
Muhammad Reza Aulia et. al. [21]	Structural analysis and symbol analysis	10000 datasets of signature images	Convolutional Neural Network	Structural Analysis:78.20-96.20% Symbol Analysis: 98.46%
Ghosh Subhankar et al [22]	Structural and Geometrical features	5300 responses from different users	Support Vector Machine and Convolutional	86.70%

			Neural Network	
Proposed Method	Local Directional Pattern, Local Phase Quantization and Shape features	4790 offline handwritten Signature with 13 different classes of signature	Random Forest Classifier	97.46% for Female 96.64% for Male

CONCLUSION

In the present research work, the personality trait analysis of a person based on Handwritten Signature is carried out. It can also allow in some aspect, disclosing the inner psychological state of person. It is observed that many researchers have worked on personality analysis of person using different methodologies and obtained promising results with own dataset and standard dataset. The main focus of the proposed work is on feature-level fusion algorithm and random forest classifier. Herein, a feature set has been prepared by extracting two local variant texture descriptors (LDP, LPQ) and Shape features from the samples in the experimental own dataset comprising an in-house total of 4790 offline handwritten signature images with 13 different classes of 250 male and 229 female signatures of good quality. With Random Forest Classifier applied to fused features,the accuracy rate of 97.64% for female signature and 96.64% for male signature is attained. In future, the personality analysis of online handwritten signatures and other biometrics using deep learning techniques will be studied.

COMPLIANCE WITH ETHICAL STANDARDS

- Conflict of Interest: The authors declare that they have no conflict of interest.
- Research involving Human and Animal Rights: This work does not contain any studies with animals performed by any of the authors.
- Ethical Standards: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institution.

ACKNOWLEDGEMENT

Authors would like to thank all students, teaching and non-teaching staff ofRani Channamma University, Belagavi, Karnataka-India, who have voluntarilygave the handwritten signature samplesfor this research work.

References

1. Mekhaznia, Tahar and Djeddi, Chawki and Sarkar, Sobhan. “Personality Traits Identification through Handwriting Analysis. (2021) 10.1007/978-3-030-71804-6_12.
2. Hemlata, & Sachan, Manoj & Singh, Shailendra. Personality Detection using Handwriting Analysis:Review. Proceedings of the 7th, International conference on ACEC-2018, ISBN: 978-1-63248-157-3, 10.15224/978-1-63248-157-3-33.
3. Nijil Raj N. , Mohammed Thaha , Sushlin Grace Shaji , Shibina S, “Forecasting Personality Based On Calligraphy Using CNN and MLP”, International Journal of Computer Sciences and Engineering, Vol.8, Issue.7, July 2020 E-ISSN: 2347-2693, DOI: https://doi.org/10.26438/ijcse/v8i7.4148 | Available online at: www.ijcseonline.org.
4. Kurowski, M.; Sroczynski, A.; Bogdanis, G.; Czyzewski, A., “An Automated Method for Biometric Handwritten Signature Authentication Employing Neural Networks. Electronics 2021, 10, 456. https:/ /doi.org/10.3390/ electronics10040456

5. Marcos Faundez-Zanuy, Julian Fierrez, Miguel A. Ferrer, Moises Diaz, Ruben Tolosana, Réjean Plamondon, "Handwriting Biometrics: Applications and Future Trends in e-Security and e-Health", Cognitive Computation (2020) 12:940–953, https://doi.org/10.1007/s12559-020-09755-z.

A. Varshney and S. Puri, "A survey on human personality identification on the basis of handwriting using ANN," 2017 International Conference on Inventive Systems and Control (ICISC), 2017, pp. 1-6, doi: 10.1109/ICISC.2017.8068634.

6. Khanam, M.. "Scientific Handwriting Changes Your Life." International Journal of Research in Engineering, Science and Management, 3(11), 13–15.(2020) Retrieved from https://www.journals.resaim.com/ijresm/article/view/363

7. Donato lmpedovo, Giuseppe Pirlo G, "Dynamic Handwriting Analysis for the Assessment of Neurodegenerative Diseases: A Pattern Recognition Perspective", IEEE REVIEWS IN BIOMEDICAL ENGINEERING, VOL. 12, 2019, Digital Object Identifier 10.1109/RBME.2018.2840679.

8. Saini, Komal & Kaur, Manpreet.. Forensic study on the effect of age and illness (Parkinsonism) on handwriting characteristics. Egyptian Journal of Forensic Sciences. (2019) 9. 10.1186/s41935-019-0131-9.

9. Taylor, Melissa K... "Forensic Handwriting Examination and Human Factors: Improving the Practice through a Systems Approach." (2020).

10. Jain, A. K., Dass, S. C., and Nandakumar, K. Can Soft Biometric Traits Assist User Recognition? Proc. SPIE - Int. Soc. Opt. Eng. 5404. (2004). Doi: 10.1117/12. 542890.

11. Singh, Harshdeep., and Verma, Priyanka.. Anita and Navjot Kaur "Analysis of Signature Features in Different Age Groups". (2019) Peer Reviewed J. Forensic Genet. Sci. 14 (2). doi:10.32474/PRJFGS.2019.03.000160.

12. Al-ma'adeed, Somaya, and HassaJne, AbdelAali. Automatic Prediction of Age, Gender, and Nationality in Offline Handwriting. EVRASIP /Image Video Process. 10. (2014) Doi: l0.1186/1687-5281-2014-10.

13. Putu Veda Andreyana , Made Sudarma , Nyoman Pramaita, "Expert System of Personality Analysis Based on Graphology Handwriting and Signature", IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 21, Issue 6, Ser. II (Nov - Dec 2019), PP 26-34 www.iosrjournals.org.

14. Maliki, I & Sidik, M.. Personality Prediction System Based on Signatures Using Machine Learning. IOP Conference Series: Materials Science and Engineering. (2020) 879. 012068. DOI:10.1088/1757-899X/879/1/012068..

15. Vaishali R. Lokhande, B. Gawali, "Analysis of signature for the prediction of personality traits", 2017 1st International Conference on Intelligent Systems and Information Management (ICISIM) DOI:10.1109/ICISIM.2017.8122145 Corpus ID: 36604659

16. Fioza Mackay Young, "Personality Analysis of your Handwritng Signature, Discover real personality behind any signature", Vol. 24, 2nd Edition.

17. Rosso OA, Ospina R, Frery AC, Classification and Verification of Handwritten Signatures with Time Causal Information Theory Quantifiers. PLoS ONE 11(12): e0166868. (2016) doi:10.1371/journal.pone.0166868.

18. Yasemin Bay Ayzeren, Meryem Erbilek , And Erbuğ Çelebi, "Emotional State Prediction From Online Handwriting and Signature Biometrics", DOI: 10.1109/ACCESS.2019.2952313, VOLUME 7, 2019.

19. Gavrilescu and Vizireanu, "Predicting the Big Five personality traits from handwriting", EURASIP Journal on Image and Video Processing (2018) 2018:57 https://doi.org/10.1186/s13640-018-0297-3,

20. Muhammad Reza Aulia, Esmeralda C. Djamal and Abdul Talib Bon, "Personality Identification Based on Handwritten Signature Using Convolutional Neural Networks", Proceedings of the 5th NA International Conference on Industrial Engineering and Operations Management Detroit, Michigan, USA, August 10 - 14, 2020.

21. Ghosh, Subhankar & Shivakumara, Palaiahnakote & Roy, Prasun & Pal, Umapada & Lu, Tong.. Graphology based Handwritten Character Analysis for Human Behavior Identification. CAAI Transactions on Intelligence Technology. 5. 10.1049/trit.2019.0051. (2020)

22. Pankaj Jalote", A Concise Introduction to Software Engineering", Springer-Verlag London DOI: https://doi.org/10.1007/978-1-84800-302-6 ,ISBN:978-1-84800-302-6, Series ISSN1863-7310,2008,

23. T. Jabid, M. H. Kabir and O. Chae, "Gender Classification Using Local Directional Pattern (LDP)," 2010 20th International Conference on Pattern Recognition, 2010, pp. 2162-2165, doi: 10.1109/ICPR.2010.373.

24. M. A. Ferrer, F. Vargas, C. M. Travieso and J. B. Alonso, "Signature verification using local directional pattern (LDP)," 44th Annual 2010 IEEE International Carnahan Conference on Security Technology, 2010, pp. 336-340, doi: 10.1109/CCST.2010.5678680.

25. Moolla, Yaseen & Viriri, Serestina & Tapamo, Jules-Raymond & Nelwamondo, Fulufhelo.. Local directional pattern based signature verification using weighted fractional distance classification. 79-83. 10.1145/2389836.2389846. (2012)

26. Heikkilä, Janne & Rahtu, Esa & Ojansivu, Ville.. Local Phase Quantization for Blur Insensitive Texture Description. Studies in Computational Intelligence. 506. 49-84. 10.1007/978-3-642-39289-4_3. (2014)

27. Gornale, Shivanand & R, Kruthi & Patil, Abhijit. Fusion of Local Binary Pattern and Local Phase Quantization features set for Gender Classification using Fingerprints. International Journal of Computer Sciences and Engineering. 7. 10.26438/ijcse/v7i1.2229. (2019).

28. Gornale S. S., Patravali P. U., Hiremath P. S... Detection of osteoarthritis in knee radiographic images using artificial neural network. Int. J. Innovat. Technol. Explor. Eng. 8, 2429–2434. 10.35940/ijitee.L3011.1081219, (2019c).

29. Gornale S. S., Patravali P. U., Manza R. R. Detection of osteoarthritis using knee X-ray image analyses: a machine vision based approach. Int. J. Comput. Appl. 145, 20–26. 10.5120/ijca2016910544. (2016b)

30. Batool, F.E., Attique, M., Sharif, M. et al. Offline signature verification system: a novel technique of fusion of GLCM and geometric features using SVM. Multimed Tools Appl (2020). https://doi.org/10.1007/s11042-020-08851-4

31. Gornale, S. S., Kumar, S., Patil, A., & Hiremath, P. S. "Behavioral Biometric Data Analysis for Gender Classification Using Feature Fusion and Machine Learning". Frontiers in robotics and AI, 8, 685966. https://doi.org/10.3389/frobt.2021.685966, (2021)

32. Shah, Abdul Salam & Shah, Masood & Fayaz, Muhammad & Wahid, Fazli & Khan, Hira & Shah, Asadullah.. Forensic Analysis of Offline Signatures Using Multilayer Perceptron and Random Forest. International Journal of Database Theory and Application. 1010. 139-148139. 10.14257/ijdta.2017.10.1.13. (2017)

33. L. Breiman "Random forests", Machine Learning, vol. 45, no. 1, (2001), pp. 5–32.