A New Swarm-Based Framework for Handwritten Authorship Identification in Forensic Document Analysis

Satrya Fajri Pratama, Azah Kamilah Muda, Yun-Huoy Choo, and Noor Azilah Muda

Computational Intelligence and Technologies (CIT) Research Group,
Center of Advanced Computing and Technologies,
Faculty of Information and Communication Technology,
Universiti Teknikal Malaysia Melaka
Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia
satrya@student.utem.edu.my,
{azah,huoy,azilah}@utem.edu.my

Abstract. Feature selection has become the focus of research area for a long time due to immense consumption of high-dimensional data. Originally, the purpose of feature selection is to select the minimally sized subset of features class distribution which is as close as possible to original class distribution. However in this chapter, feature selection is used to obtain the unique individual significant features which are proven very important in handwriting analysis of Writer Identification domain. Writer Identification is one of the areas in pattern recognition that have created a center of attention by many researchers to work in due to the extensive exchange of paper documents. Its principal point is in forensics and biometric application as such the writing style can be used as bio-metric features for authenticating the identity of a writer. Handwriting style is a personal to individual and it is implicitly represented by unique individual significant features that are hidden in individual's handwriting. These unique features can be used to identify the handwritten authorship accordingly. The use of feature selection as one of the important machine learning task is often disregarded in Writer Identification domain, with only a handful of studies implemented feature selection phase. The key concern in Writer Identification is in acquiring the features reflecting the author of handwriting. Thus, it is an open question whether the extracted features are optimal or near-optimal to identify the author. Therefore, feature extraction and selection of the unique individual significant features are very important in order to identify the writer, moreover to improve the classification accuracy. It relates to invarianceness of authorship where invarianceness between features for intra-class (same writer) is lower than inter-class (different writer). Many researches have been done to develop algorithms for extracting good features that can reflect the authorship with good performance. This chapter instead focuses on identifying the unique individual significant features of word shape by using feature selection method prior the identification task. In this chapter, feature selection is explored in order to find the most unique individual significant features which are the unique features of individual's writing. This chapter focuses on the integration of Swarm Optimized and Computationally Inexpensive Floating Selection (SOCIFS) feature selection technique into the proposed hybrid of Writer Identification framework

A.K. Muda et al. (eds.), *Computational Intelligence in Digital Forensics*,
Studies in Computational Intelligence 555,
DOI: 10.1007/ 978-3-319-05885-6_16, © Springer International Publishing Switzerland 2014

and feature selection framework, namely Cheap Computational Cost Class-Specific Swarm Sequential Selection (C_4S_4). Experiments conducted to proof the validity and feasibility of the proposed framework using dataset from IAM Database by comparing the proposed framework to the existing Writer Identification framework and various feature selection techniques and frameworks yield satisfactory results. The results show the proposed framework produces the best result with 99.35% classification accuracy. The promising outcomes are opening the gate to future explorations in Writer Identification domain specifically and other domains generally.

Keywords: swarm-based framework, feature selection, handwritten authorship, significant features, forensic document analysis.

1 Introduction

Everyone in this world possesses their own uniqueness, whether in physical, appearance, and characteristics. These unique features are making each and every person discernible from the others. Generally, unique features used to identify an individual are biological feature, such as fingerprint, handprint, hand geometry, face, or voice. There is one feature which is not commonly used, even not a part of biological feature, which is handwriting [1]. This feature is a derivate feature of hand geometry, but also affected by other factors. The complexities of the process to produce handwriting, even the simplest alphabet letter, making this process is capable to identify someone. Even when two writers produce two handwritings that look similar, there are some features that can be used to differentiate their writings. Meaning, even someone can fake the handwriting of another person, but there are some features exist only in the original writing, this is because the original and the fake writings are having different features. Even though in the reality the handwriting will be changed due to its writer's physical and emotional condition, the unique features of one person always exist on his writing, regardless of the condition. Due to its uniqueness and consistency, the features in the handwriting are used to analyze and authenticate forensics documents [2].

The use of handwritten paper documents has never been diminished although the world has lived in digital age for quite some time. There have always been situations in which unsigned or anonymous writings on documents were potentially important. Thus, the provision of proof respecting the authorship of such documents has long been an issue [2]. The Questioned Document Examination (QDE) is an area of the Forensic Science with the main purpose to answer questions related to questioned document (authenticity, authorship and others) and has a large field of applications. There are basically two different sub-areas in the QDE: the document analysis and the handwriting analysis [2]. The first one evaluates the structural analysis of the document to find adulteration, falsification, obliteration and others, while the second investigates the originality or the association between one or more manuscripts to an author [3], for instance when validating the purchase using credit cards, where the card's owner signature on the receipt is slightly different than the signature stored by the bank, or in the opposite situation where the forger signature is similar to the card's owner. Handwriting analysis is applied to many types of investigation like fraud, homicide, suicide and others, and it has two basic analysis subjects, manuscripts and signatures. Even with distinct features, both keep a narrow relation having the same root or origin in the writer's learning process, in other words, they carry the experiences acquired by the writer during and after his learning process through the improvement of the handwriting personal style [2].

The handwriting analysis research field consists of two categories, which are handwriting recognition and handwriting identification. Fig. 1 depicts the handwriting analysis domain. Handwriting recognition deals with the contents conveyed by the handwritten word, while handwriting identification tries to differentiate handwritings to determine the author [4]. Handwriting identification can be categorized into handwritten authorship identification, handwritten authorship characterization, and similarity detection. Authorship characterization is aimed at inferring an author's background characteristics rather than identity. Similarity detection compares multiple pieces of writing without identifying the author. Handwritten authorship identification, or simply known as authorship identification, evaluates the possibility of one author produces a written document by examining other documents produced by that author [5]. Although authorship identification is categorized as QDE research area, it has evolved into its own matured domain, where the application of authorship identification is not always related to QDE. Authorship identification contributes great importance towards the criminal justice system and has been widely explored in forensic handwriting analysis [4, 6-12]. Nevertheless, there are also many issues and scenarios in authorship identification that pose as challenges which require further investigations and explorations.

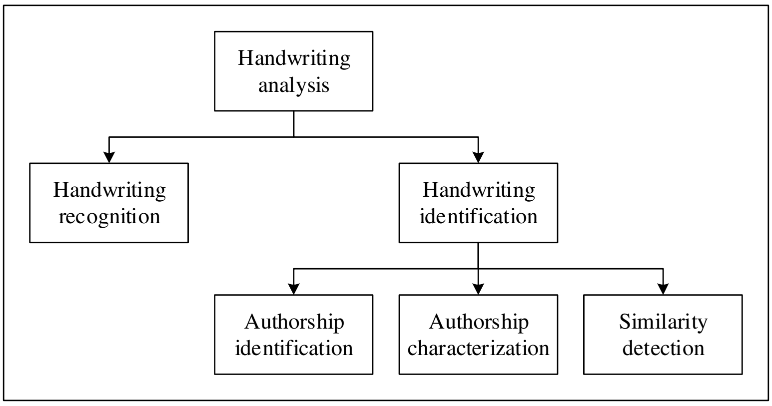


Fig. 1. Handwriting analysis domain [5]

The performance of pattern recognition applications is heavily depended on the feature extraction and classification method employed [13, 14], which leads to the key concern issue in authorship identification: acquiring the features reflecting the author of handwriting, namely unique individual significant features [4, 11, 15-20].

The essence of authorship identification is to identify a set of features that remain relatively constant among a number of writings by a particular author, and in such a process, the classification technique is very important to the performance of authorship identification [5]. A survey conducted by [5] found a number of studies that show the discriminating power of different types of features, by which researchers attempt to identify an optimal set of features for authorship identification. There are several broad categories for authorship identification, which are platform, author resolvability, text dependency, and individuality of handwriting [4, 21, 22], and shown in Fig. 2.

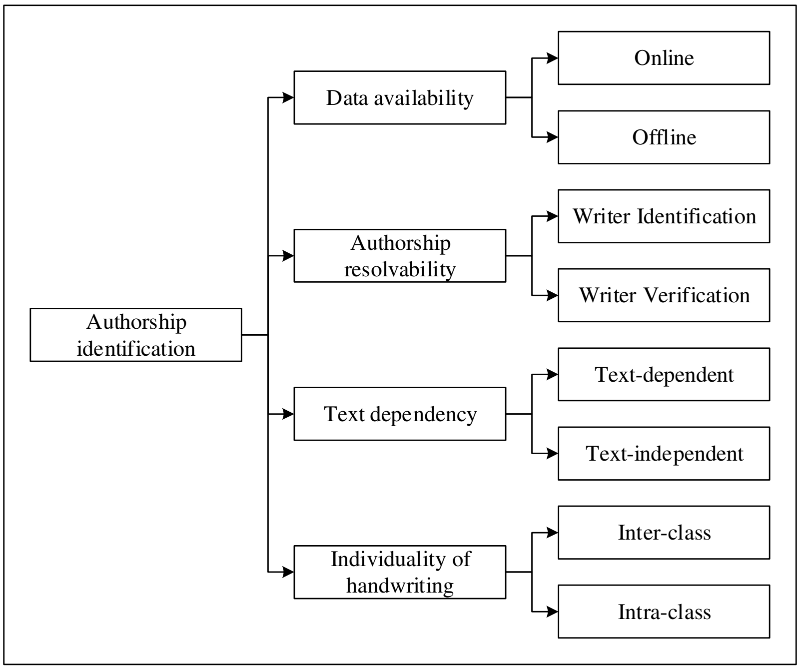


Fig. 2. Authorship identification category [4, 21, 22]

The first category of authorship identification is the platform of the system itself. The platform of the system can be categorized into two, which are offline system and online system [4]. The terms of offline and online system are referring to the input method of the system, rather than the location of the system (as the web application or stand-alone desktop application). Offline system acquires its input from scanned documents or images, while online system acquires its input from touch-sensitive, motion-sensitive, gesture-sensitive, and pressure-sensitive acquiring devices, such as tablets, and thus contains temporal information and theoretically should provide more accurate results [4, 21]. Therefore, online and offline systems have different set of problems and information and thus require different processing methods.

The second category is author resolvability, which consists of two domains: Writer Identification (WI) and Writer Verification (WV). WI performs a one-to-many search in a large database with handwriting samples of known authorship and returns a likely list of candidates, while WV involves a one-to-one comparison with a decision whether or not the two samples are written by the same person, by determining whether the distance between two chosen samples is smaller than a predefined threshold [22]. Furthermore, there are two modes of WV, claim verification and questioned document verification. In the first mode, the system verifies the claim made by a person previously enrolled in the system, while in the second mode, verification problem verifies whether two given documents, questioned document, whose identity need to be verified and reference document, which is collected from the writer for comparison, belong to the same writer or not. The writer of the reference document may or may not be known. The difference between the two is that in this case no database of writers is available and thus, a threshold cannot be computed. In order to solve the problem, some statistical measure such as such as hypothesis testing, standard deviation, and mean square error is needed to compute the significance of the score [21, 22].

On the other hand, WI can be included as a particular kind of dynamic biometric in pattern recognition for forensic application. WI distinguishes writers based on the shape or individual writing style while ignoring the meaning of the word or character written, due to the differences between one author to another in terms of character association, shape, and the writing style [4, 9, 11, 23-26]. Although there are variances of writing in times, the individual writing style is persistent [4, 9, 11, 23, 27, 28]. And thus, the significant individual features are generalized as the unique features that are persistent regardless of the handwriting shape. The key concern in WI is in acquiring the features reflecting the author of handwriting [4, 11, 15-20]. Thus, it is an open question whether the extracted features are optimal or near-optimal to identify the author. [29] discussed several experiments conducted by various researchers in order to improve WI. [30] treated WI as a texture analysis problem using multichannel Gabor filtering and grey-scale co-occurrence matrix techniques, [31] and [32] addressed the problem of writer verification by casting it as a classification problem with two classes: authorship and non-authorship, [33] morphologically processed horizontal projection profiles on single words, [34] and [35] proposed edge-based directional probability distributions and connected component contours as features, [36] introduced graphemes as features for describing the individual properties of handwriting, and [37] presented a set of eleven features which can be extracted easily and used for the identification and verification of documents containing handwritten digits.

From text dependency point of view, authorship identification can be divided into two broad categories, which are text-dependent and text-independent methods. The text-dependent methods are very similar to signature verification techniques and use the comparison between individual characters or words of known semantic content, and therefore require the prior localization and segmentation of the relevant information. The text-independent methods use statistical features extracted from the entire image of a text block, and thus a minimal amount of handwriting is necessary in order to derive stable features insensitive to the text content of the samples [4, 22]. Text-dependent methods provide high accuracy and confidence with small amount of data, which is practically not possible for text-independent systems. However, they are more prone to forgery, as the verification text is known in advance. In case of text-independent systems, forgery is not a major problem as the text-independent systems extract less frequent properties from the handwritten document that are difficult to forge [4, 21, 38].

The last category, individuality of handwriting is deemed as the most important issue in authorship identification, which is the main key to identify the author and is closely related to feature extraction task, and thus it is defined as the variance between features for intra-class must be lower than variance between features for inter-class [4]. It relies on two principles: (1) habituation, since people are primarily creatures of habits and writing is the collection of those habits, which are considered neither instinctive nor hereditary but are complex processes that are developed gradually, and (2) individuality or heterogeneity of handwriting, in which each individual had his own style of writing and no two individuals can have the same handwriting [21]. It is only possible to the extent that the variation in handwriting style between different writers exceeds the variations intrinsic to every single writer considered in isolation [22]. It can be proven using similarity error [25, 33, 37, 39] and has been explored by many researchers [4, 26, 28, 39].

In theory, the discriminating power directly relates to the number of features, nevertheless the vast machine learning algorithms practical experiences often proves this does not always apply. The learning process becomes more and more difficult during the training phase if there are too many irrelevant and redundant information, or worse, if the data is noisy and unreliable [40, 41]. Coherent with this traditional concept, the search for the unique feature for every individual in WI domain must consider the condition where the feature for one author may be similar to other authors, and thus should be omitted because of its non-uniqueness. This search objective is similar to the purpose of the feature selection, where the resulting subset is the discriminator between one classes to other classes. Hence, the feature selection phase should be incorporated after feature extraction phase in WI framework, and thus reduce the number of features used and improve the classification performance and accuracy [42]. Since features are regarded as an abstract representation of handwriting, the quality of the feature selection directly influences this representation [5]. Therefore, the purpose of feature selection in this chapter is to acquire the unique features that represent the author of the handwriting in WI domain.

Many previous works have explored the use of feature selection in WI domain [5, 29, 37, 43, 44]. And yet, these studies have not fully addressed the issue in WI domain itself, because instead of acquiring the unique individual significant features to reflect the author of handwriting, these studies focus on the acquiring the features that distinguish one author to another. While the general and common approach does produce good result, it has no significant differences with other pattern recognition problems, since the concept of Individuality of Handwriting is not apparent. Individuality of Handwriting is the most important issue in WI domain, which is the main key to identify the handwritten authorship and is closely related to feature extraction task.

Motivated by the success of the framework proposed by [4], where the global features of handwriting is extracted and thus the Individuality of Handwriting is preserved by using Invariant Discretization, this chapter is trying to further improve the quality of the global features of handwriting produced by acquiring the features that is representing the author of the handwriting using feature selection technique. These representative features must always be existed in every handwriting produced by the same author and should provide enough discriminating power to differentiate the author from other authors. These discriminative features are called unique individual significant features. Because the unique individual significant features are different from one author to other authors, general pattern recognition framework may not be suitable for acquiring these features. The framework employed for this specific task must be capable of acquiring different set of significant features for every author. There are several existing frameworks that is capable of acquiring class-specific features subset, however these frameworks should be modified prominently or they employs feature selection technique that is not suitable for acquiring unique individual significant features.

Therefore, a robust framework to cater this problem must be developed, and at the same time, employs the effectiveness of feature selection to acquire the unique individual significant features. Embarking from these motivations, this chapter is conducted in order to devise a novel feature selection technique which is capable to acquire these unique individual significant features. Furthermore, the proposed technique itself is not working on its own. The proposed technique is developed as a part of vigorous framework, specifically devised for WI domain. The proposed framework employs proposed feature selection technique as the mechanism to acquire the unique individual significant features which is unique to each author. The acquisition of unique significant features also allows the performance of the proposed framework to exceed the performance of existing WI framework [4].

2 Existing frameworks for Handwritten Authorship Identification in Forensic Document Analysis

Writer Identification (WI) is an active area of research in pattern recognition due to extensive exchange of paper documents, although currently the world has already moved toward the use of digital documents. WI distinguishes writers based on the handwriting, and ignoring the meaning of the words. Previous studies have explored various methods to improve WI domain, and these studies produced the satisfying performance. However, the use of feature selection as one of important machine learning task is often disregarded in WI domain, which has been proven in the literature where only a handful of studies implemented feature selection task in the WI domain [29, 43, 44].

The key concern in WI is in acquiring the features reflecting the author of handwriting. Although WI is still attracting a vast array of researches since a long time, predominantly in forensic and biometric applications, the question of whether the extracted features are optimal or near-optimal to identify the author is still remain unanswered. This is because the extracted features may include many garbage features. Such features are not only useless in classification, but sometimes degrade the performance of a classifier designed on a basis of a finite number of training samples [4, 42, 45-49]. The features may not be independent of each other or even redundant. Moreover, there may be features that do not provide any useful information for the task of WI [29, 41, 42]. Therefore, feature extraction and selection of the unique individual significant features are very important in order to identify the writer, moreover to improve the classification accuracy.

Handwritten words are very effective in discriminating handwriting, and thus in the study conducted by [4], the holistic approach of global features is used where cursive word is defined as one indivisible entity and extracted by using United Moment Invariant (UMI) [50] technique. Individual features can be acquired by using feature selection technique, by selecting the subset of features. Although in theory, more features provide more discerning power, but in the reality it will degrade significantly the performance [40]. Thus, it is vital to acquire individual features and to perform feature selection for these features, because this will provide simpler identification process and improve the performance of identification in identifying the author.

WI is a part of pattern recognition domain, specifically in handwriting analysis. Thus, traditional pattern recognition framework is appropriate for solving the problem of WI, which is pre-processing, feature extraction and classification. The most recent work to enhance the traditional WI framework is the introduction of an enhanced framework specifically for WI domain proposed by [4], termed as Enhanced WI Framework (EWIF), which consists of feature extraction, feature discretization, and classification. The framework design for traditional pattern recognition framework and EWIF are shown in Fig. 3 and Fig. 4 respectively.

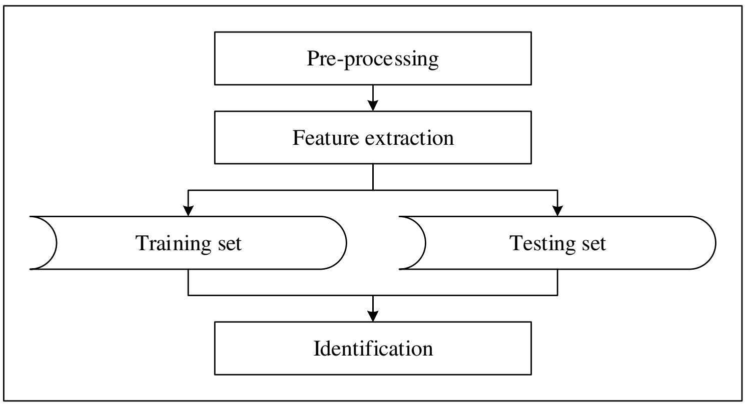


Fig. 3. Traditional pattern recognition framework

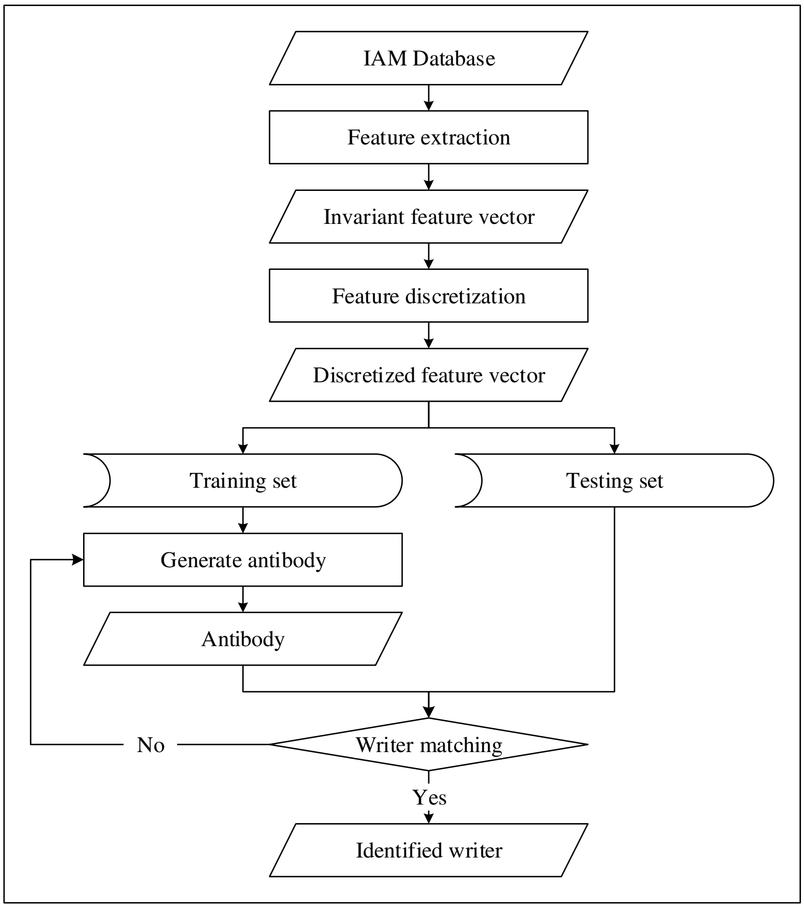


Fig. 4. Enhanced WI Framework [4]

Feature extraction is a process of converting input object into feature vectors. The extracted features are in real value and unique for each word. By using UMI, a digital image is converted to a set of moments which represents the global characteristics of an image shape. Global Moment Function can be used to generate a set of moments that uniquely represent the global characteristic of an image. Moments are scalar quantities used to characterize a function and to capture its significant features. Moment Invariants are very useful tools for pattern recognition [50]. The first introduction of Moment Invariants to pattern recognition and image processing was the employment of algebraic invariants theory by [51], which derived his renowned seven invariants to the rotation of 2D objects. And thus ever since, it has been chosen as one of the most important and frequently used shape descriptors options. Even though they suffer from certain intrinsic limitations (the worst of which is their globalness, which prevents direct utilization for occluded object recognition), they frequently serve as "first-choice descriptors" and as a reference method for evaluating the performance of other shape descriptors [52]. Geometric Moment Invariants (GMI) [51] presents a set of moments based on combinations of algebraic invariants. This is complied with the definition of invariants given by [53]: an image or a shape feature is invariant if that image or shape undergoes one or a combination of linear transformations. The moments are normalized using (1).

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{(p+q+2)/2}} \tag{1}$$

where μ_{pq} is the first, second, and third order of moment which represent the center of the image, measure the variance of the image intensity distribution, and denotes the projection of the image respectively, μ_{00} is the zero-th order moment which represents the total intensity of the image, and $p + q = 2, 3, 4, \ldots$. These moments are invariant under the image scale, translation and rotation, and thus there are seven tuples of moment invariant proposed, which are shown in (2).

$$\begin{aligned}
&\phi_1 = \eta_{20} + \eta_{02} \\
&\phi_2 = \phi_1^{\,2} + 4\eta_{11}^{\,2} \\
&\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (\eta_{03} - 3\eta_{21})^2 \\
&\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\
&\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})\left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2\right] \\
&+ (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})\left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2\right] \\
&\phi_6 = (\eta_{20} - \eta_{02})\left[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2\right] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\
&\phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})\left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2\right] \\
&+ (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})\left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2\right]
\end{aligned} \tag{2}$$

However, [54] found that GMI lose its scale invariance in discrete condition. Several improvements to maintain scale invariance are made by [55-57]. All these improvements are not valuable based on both regions and boundaries simultaneously or the formulas are not coincident with Hu's moments. Therefore, [50] proposed new Moment Invariants called United Moment Invariants (UMI), which is capable of keeping invariant to region and closed and unclosed boundary, both in discrete and continuous condition. The equation of UMI is as shown in (3).

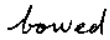
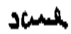
$$
\begin{aligned}
\theta_1 &= \frac{\sqrt{\phi_2}}{\phi_1} \\
\theta_2 &= \frac{\phi_6}{\phi_1\phi_4} \\
\theta_3 &= \frac{\sqrt{\phi_4}}{\phi_4} \\
\theta_4 &= \frac{\phi_5}{\phi_3\phi_4} \\
\theta_5 &= \frac{\phi_1\phi_6}{\phi_2\phi_3} \\
\theta_6 &= \left(\phi_1 + \sqrt{\phi_2}\right)\frac{\phi_3}{\phi_6} \\
\theta_7 &= \frac{\phi_1\phi_5}{\phi_3\phi_6} \\
\theta_8 &= \frac{\phi_3 + \phi_4}{\sqrt{\phi_5}}
\end{aligned}
\tag{3}
$$

where ϕi are GMI. The features extracted by UMI are the pattern to represent the image shape. It is also worth mentioning that [50] also found the scale invariance of GMI is untenable in discrete condition and the disunion of invariants formula based on region and boundary. The information of different types of geometrical features of the image is also provided by UMI [58]. The feature extraction phase in this chapter is achieved by using global representation of UMI [50] to acquire the global features of handwriting image, due to the requirement of cursive word is needed to extract as one single indivisible entity.

According to [4], the advantages of global approach are including its capabilities to show the individuality of handwriting [23], is shown to be very effective in reducing the complexity of the word [59], moreover to increase the accuracy of classification, and it is invariant with respect to all different writing styles; hence it holds immense promise for realizing near-human performance [60] and very robust in detecting similar object when it is used in similarity search. Table 1 is the example of feature invariant of words using UMI with eight features vector for each image, with f1 represents the first feature, f2 for second feature, and henceforth.

Many real-world classification tasks exist that involve continuous features where such algorithms could not be applied unless the continuous features are first discretized. Discretization is a process of dividing a range of continuous features into disjoint intervals, which labels can then be used to replace the actual data values [61]. Discretization engages searching for cut-off points that determine intervals and thus unifying the values over each interval. All values that lie within an interval are

Table 1. UMI representation for handwritten word image

Word	f1	f2	f3	f4	f5	f6	f7	f8
	1.84	1.79	0.91	1.31	0.84	1.00	0.73	1.79
	1.53	1.08	1.12	1.96	0.72	1.49	1.82	1.46
	1.61	1.53	0.53	0.38	0.80	1.26	0.25	3.29
	1.99	8.24	0.65	0.76	3.77	0.20	0.09	2.40
	3.08	2.06	0.52	0.64	0.52	0.82	0.31	2.75

mapped to the same value, in effect converting numerical attributes that can be treated as being symbolic [62]. Discretization largely contributes to rough set theory [63] and provides more comprehensible knowledge representation which is reduced and simplified [64], and thus more accurate and faster. Continuous variable discretization has recently received significant attention in the machine learning domain [65]. The goal of discretization is to find a set of cut points to partition the range into a small number of intervals that have good class coherence, which is usually measured by an evaluation function. In addition to the maximization of interdependence between class labels and attribute values, an ideal discretization method should have a secondary goal to minimize the number of intervals without significant loss of class-attribute mutual dependence [66].

Discretization is usually performed prior to the learning process and it can be broken into two tasks. The first task is to find the number of discrete intervals. Only a few discretization algorithms perform this; often, the user must specify the number of intervals or provide a heuristic rule. The second task is to find the width, or the boundaries, of the intervals given the range of values of a continuous attribute [66]. Discretization is in EWIF employed because the features extracted from feature extraction phase are in continuous forms [61]. [4] argues that discretization is important in order to obtain the detachment of writers' individuality and produce better data representation, thus [4] proposes a new discretization technique namely Invariant Discretization, which provides standard representation of individual features, which allows small variance between features for intra-class (same author) and large variance for inter-class (different authors). Although feature discretization provides better representation of individual features, this mechanism only partially reflect the key concern in WI domain.

In the classification phase, the correct identification accuracy, or termed classification accuracy is calculated to measure the quality of feature subset produced from feature selection phase. EWIF applies Modified Immune Classifier (MIC) [4]. In order to detect the same features of a writer, the detector must complement the antigen. The detector in this case is the features extracted from training dataset, while the antigen is the features extracted from testing dataset or questioned handwriting [4]. The classifier consists of two modules: censoring module and monitoring module. The detector is generated with the complementary of the self-cell in the censoring module.

On the other hand, monitoring module is the matching process, and the term "bind" is adopted in order to describe the matching process, which is due to the complementary of the self-cell which is defined as the detector in the censoring module [4]. MIC uses several binary matching techniques, which are Hamming distance, *r*-Chunk, *r*-Contiguous, and Multiple *r*-Contiguous. The binary strings are used to represent the detectors and antigens, which forms the binary matching rule.

The inclusion of feature selection calls for the further improvement to the EWIF. This is because the main drawback of EWIF is that the mechanism to acquire the unique significant features is not present and is not defined as the part of the framework; instead the whole features are used for the identification phase. The acquisition of the unique significant features is apparently one of the important issues on WI domain because it provides more effective way to identify the handwritten authorship [4], and this issue is not addressed in the EWIF.

3 Swarm-Based Feature Selection Technique

Feature selection has become an active research area for decades, and has been proven in both theory and practice [40]. The main objective of feature selection is to select the minimally sized subset of features as long as the classification accuracy does not significantly decreased and the result of the selected features class distribution is as close as possible to original class distribution [42].

The feature set produced from feature extraction phase in traditional framework or discretization phase in Enhanced Writer Identification Framework (EWIF), may consist of relevant and irrelevant features. There will be more complexities produced in terms of accuracy and performance, if these features are used directly in classification phase. Although in theory, more features provide more discerning power, but in the reality it will degrade significantly the performance [40]. Hence, the feature selection phase should be incorporated after discretization phase, and thus reduce the number of features used and improve the classification performance and accuracy [42]. Feature selection phase should be able to filter those features and select the most unique individual significant features in the process. Therefore, selection of the unique individual significant features is very important in order to identify the writer.

Wrapper feature selection method possesses unique advantages and disadvantages. A wrapper algorithm explores the space of features subsets to optimize the induction algorithm that uses the subset for classification. The rationale for wrapper methods is that the induction method that will ultimately use the feature subset should provide a better estimate of accuracy than a separate measure that has an entirely different inductive bias [41, 67]. These methods based on penalization face a combinatorial challenge when the set of variables has no specific order and when the search must be done over its subsets since many problems related to feature extraction have been shown to be NP-hard [68]. Advantages of wrapper method are the ability to include the interaction between feature subset search and model selection, and take into account feature dependencies. On the other hand, the disadvantages are that it has higher risk of over-fitting than filter methods and are very computationally intensive [69].

Therefore, this section describes the method to optimize selected feature selection technique, particularly in diminishing computational cost.

Several techniques have been introduced throughout the last decade to reduce the complexity of wrapper method, for instance is by infusing it with recent stochastic optimization [44, 70-75], controlling the number of cross-validation [76], and hybridizing with filter methods [77, 78]. However, very few studies conducted in utilizing concurrent programming techniques [79-81]. Studies shown that implementing concurrent programming, specifically multithreading, sanctions much lesser processing time [79-83]. Therefore, the first optimization applied towards wrapper method is multithreading. This decision is motivated by the fact that wrapper technique is computationally expensive; therefore it constrained the possibility of hybridization since it will consume more resources and requires higher computational cost, and hence direct hybridization with stochastic optimization may not be the wisest option.

Considering the advantages of switching from sequential programming towards concurrent programming, or in this case is multithreading, and the lack of focus for multithreading in feature selection techniques, leads to the decision to adapt multithreading in Sequential Forward Floating Selection (SFFS) [84]. SFFS is an extension of Sequential Forward Selection [85], which suffers from the nesting effect, meaning that once a feature is included in some step of the iterative process, it cannot be excluded in a later step. SFFS performs a simple hill-climbing search. The best feature subset S is initialized as the empty set and perform the forward selection, where in each step a new subset is generated first by adding a feature x^+, but after that features x^- is searched for to be eliminated from S until the classification accuracy $J(S \setminus x^-)$ decreases, which is called as backward selection. The iterations continue until no new feature can be added because the classification accuracy $J(S \cup x^+)$ does not increase.

Multithreaded SFFS is capable to reduce the computational cost of original SFFS, not only because of the introduction of multithreading, but also because of the introduction of a novel mechanism called merit pooling. Merit pooling refers to the process of pre-calculating and storing the merit of each feature before the selection process take hand. This mechanism reduce a great deal of processing time, because instead of recalculating the merit of the feature subset every time a feature is added or removed, the proposed technique will simply sum up the merit values for each individual feature in the subset which has been stored in the merit pool previously. The resulting merit value will also be stored in the merit pool, so that future subset that has same feature member will simply use this value, without having to re-looking up the merit of individual member in the merit pool. In the original implementation of SFFS, each time a feature is added or removed from the feature subset, the merit of the subset will be calculated by repeatedly calling the induction algorithm. The process of calculating the merit is oftentimes the primary source of high computational cost of wrapper methods [41].

However, it is found that multithreaded SFFS performs not as well as original SFFS, although it opens the possibility of hybridization with swarm intelligence. In this chapter, Particle Swarm Optimization (PSO) [86, 87] is selected as the best way to optimize multithreaded SFFS. PSO is a population-based optimization method, which can be used to solve a wide array of different optimization problems. PSO is a stochastic algorithm that does not need gradient information derived from the error function. This allows the PSO to be used on functions where the gradient is either unavailable or computationally expensive to obtain. The origin of the PSO was based on the sociological behavior of bird flocking [87]. PSO initially identifies some particle as the best particle in a neighborhood of particles based on its fitness. All the particles are then accelerated in the direction of this particle, but also in the direction of their own best solutions that they have discovered previously. All particles also have the opportunity to discover better particles, in which case the other particles will change direction and head towards the new "best" particle. By approaching the current best solution, the neighboring solutions will be discovered by some of the particles. It is important to realize that the velocity term models the rate of change in the position of the particle.

The success of PSO implementation on the Writer Identification (WI) domain has also been demonstrated by [88]. Other consideration taken for selecting PSO is also due to its simple yet effective implementation. Because of this characteristic, PSO is not increasing the computational complexity of multithreaded SFFS more than necessary. The hybridization with PSO is primarily to prevent the multithreaded SFFS selects the local optima, and forces it to reevaluate the candidates with the same merit in every iteration to find the global optima. The hybrid between two techniques is dubbed Swarm Optimized and Computationally Inexpensive Floating Selection (SOCIFS). The main idea of SOCIFS is that fitness function of PSO is modified, by implementing the classification accuracy of unique individual significant features acquired by using multithreaded SFFS. This is to allow the most optimal interaction between PSO and multithreaded SFFS, and thus allow for wider search space exploration. Furthermore, there are multiple instances of multithreaded SFFS executed concurrently; each of it is executed in PSO particle. Fitness function $f(X(t))$ in SOCIFS is defined in (4) and (5), derived from [87].

$$X = \begin{cases} S \cup x^{+}, \text{if forward selection} \\ S \setminus x^{-}, \text{if backward selection} \end{cases} \tag{4}$$

$$f(X(t)) = \alpha \times \gamma_{X(t)} + \beta \times \frac{|N| - |X(t)|}{|N|} \tag{5}$$

where $\gamma_X(t)$ is the merit of particle i current subset X in iteration t, where the value is obtained by multithreaded SFFS. $|N|$ is the number of features, while $|X(t)|$ is the size of selected feature subset. α and β are the parameters used to determine the importance of classification accuracy and the subset size, where $\alpha \in [0, 1]$ and $\beta = 1 - \alpha$. Each particle will examine different feature subset and thus produce unique results, this is because the examined feature subset and its results are recorded, to prevent different particles examine the same subset multiple times. The algorithm of SOCIFS is illustrated in Fig. 5.

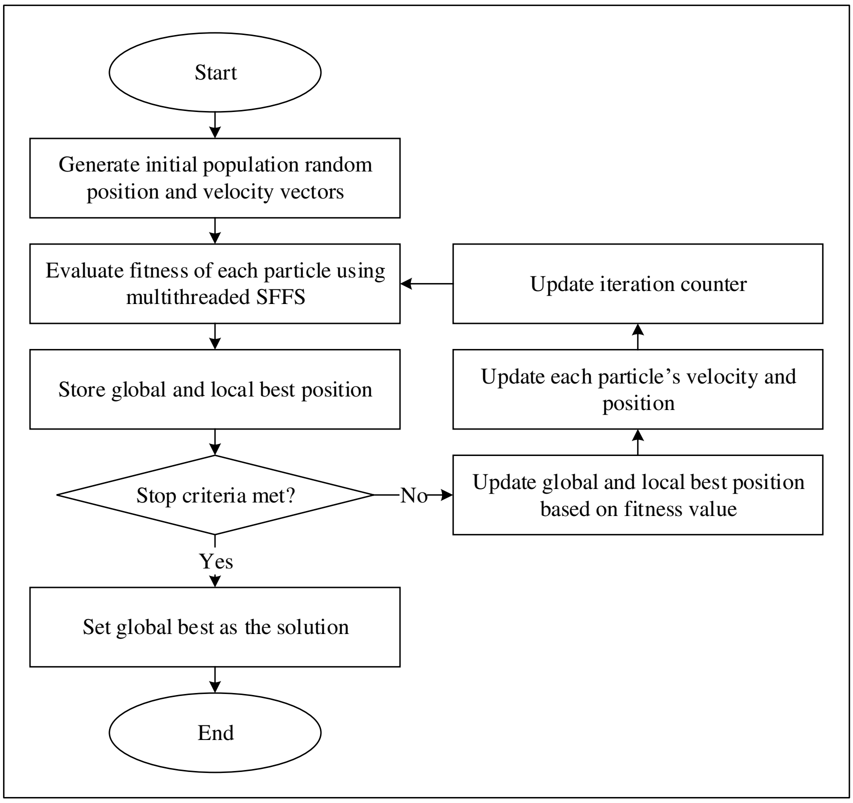


Fig. 5. Swarm Optimized and Computationally Inexpensive Floating Selection (SOCIFS)

4 Swarm-Based Framework for Handwritten Authorship Identification

The main issue in Writer Identification (WI) is to acquire the individual features from various handwritings [4]. Among these features are exists the significant individual features which directly unique to those individual. Based on this description, it is concluded that each individual possess different unique significant feature. Therefore, class-specific feature selection must be incorporated in order to capture these unique individual significant features. Even though traditional feature selection techniques can be used for acquiring these unique individual significant features [89-92], it may not be appropriate and feasible. And thus, the traditional handwriting identification framework, which consists of pre-processing, feature extraction and classification [93] is not adequate for this issue. Enhanced WI Framework (EWIF) shown in Fig. 4 [4], consists of feature extraction, feature discretization, and identification has been adopted by [90, 92] and produced good result, and therefore it can be concluded that this framework is can be further improved.

This section describes proposed swarm-based framework to cater with this class-specific feature selection issue, namely Cheap Computational Cost Class-Specific Swarm Sequential Selection (C_4S_4). Furthermore, the proposed framework is similar with General Framework for Class-Specific (GFCS) feature selection framework [94], which is shown in Fig. 6. Therefore, it can be assumed the proposed framework is a hybrid of GFCS and EWIF. And thus, several modifications should be implemented in GFCS, considering that GFCS is proposed to handle wide-range of application and domain. The proposed framework differs from GFCS and EWIF in several aspects. The differences between these frameworks are summarized in Table 2.

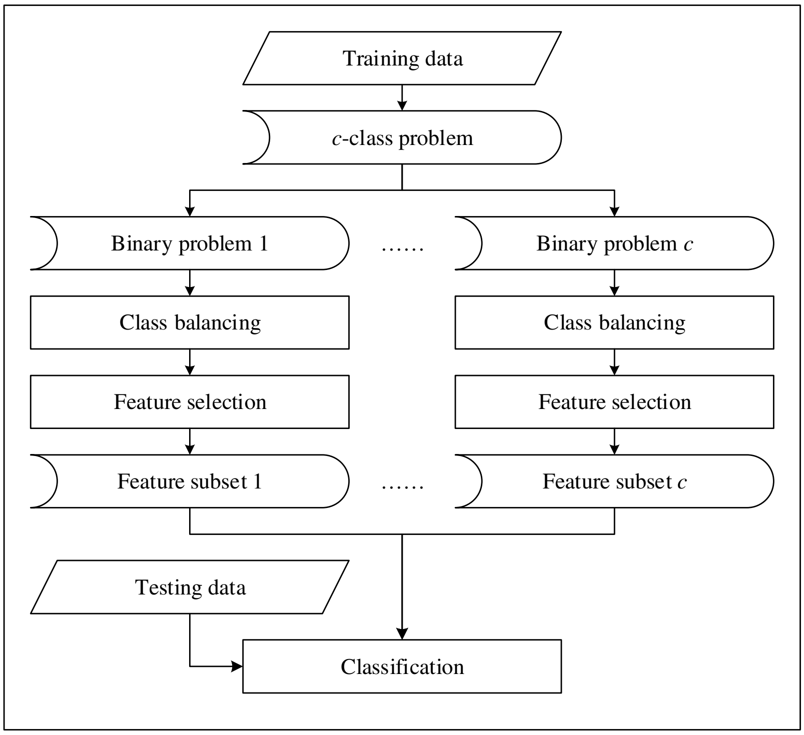


Fig. 6. General Framework for Class-Specific Feature Selection (GFCS) [94]

Table 2. Summary of EWIF, GFCS, and C_4S_4 differences

Criteria	**EWIF**	**GFCS**	**C_4S_4**
Feature extraction	Yes	-	Yes
Feature discretization	Yes	-	Yes
Class binarization	-	Yes	Yes
Class balancing	-	Yes	Yes
Feature selection	-	Yes	Yes
Antibody pool	Yes	-	Yes

GFSC is selected in this as the basis for the proposed feature selection framework because it is designed to select the class-specific feature subset, which is similar to the concept of acquiring the unique significant features in WI domain. The first difference of C_4S_4 to GFCS is that the C_4S_4 includes feature extraction and feature discretization stage, originating from EWIF, and thus produced training and testing set. After that, the framework works similarly with GFCS, which is to use the one-against-all class binarization in order to transform a *c*-class problem into *c* binary problems. For each class w_i, $i = 1,\ldots,c$; a binary problem $\langle w_i \Omega_i \rangle$ where $\Omega_i = \bigcup_{j=1, j\neq i}^{c} w_j$, is created for the training data. For each binary problem the instances of the class w_i are used as positive examples, and the instances of all other classes are used as negative examples. The generated binary problems could be imbalanced; therefore the next stage is necessary to balance the classes by applying an oversampling by repeating training instances method. $\beta_i = |w_i| - |\Omega_i|$ is then computed in the next stage, where $|w_i|$ is the number of instances in class w_i, and $|\Omega_i|$ is the number of instances in the remaining classes. If $\beta_i > 0$, the classes will be balanced by repeating instances in the class w_i until the number of instances in w_i and Ω_i are the same. For each binary problem, features are selected in the third stage by using Swarm Optimized and Computationally Inexpensive Floating Selection (SOCIFS), and the selected features are assigned to the class from which the binary problem was constructed. In this way, *c* possible different feature subsets are obtained, one for each class of the original *c*-class supervised classification problem, or unique individual significant features in this domain. These *c*-feature subsets are in turn is transformed into *c*-antibodies and stored in antibody pool that consists of all antibodies, which in turn is used in identification stage.

On the other hand, the first difference between C_4S_4 and EWIF is that the feature discretization is conducted before splitting dataset into training and testing dataset in EWIF, whereas the feature discretization is conducted after the dataset has been split into training and testing dataset in C_4S_4. This process is closely representing the real-life applications, where the testing dataset is not available to the system beforehand and thus should not be included in the training process. However, this process aroused another problem, since the training dataset is discretized while the testing dataset is not. The same discretization method cannot be directly applied to the testing dataset, because it will produce different set of data due to different cut-off points and intervals is employed. This problem is solved in C_4S_4 by storing the discretization rules for each class, which are the cut-off points and intervals. These discretization rules are employed during the classification phase, where the testing data will be discretized using each class discretization rule before the matching process is performed. In another word, an instance of the testing data will be casted into c number of instance with different values due to different rule before the identification phase will take place. The classification results for these c-instances of testing data will be ranked. The class that corresponds to the discretization rule with the highest ranking will be identified as the final class. These processes are the framework of C_4S_4, which is illustrated in Fig. 7.

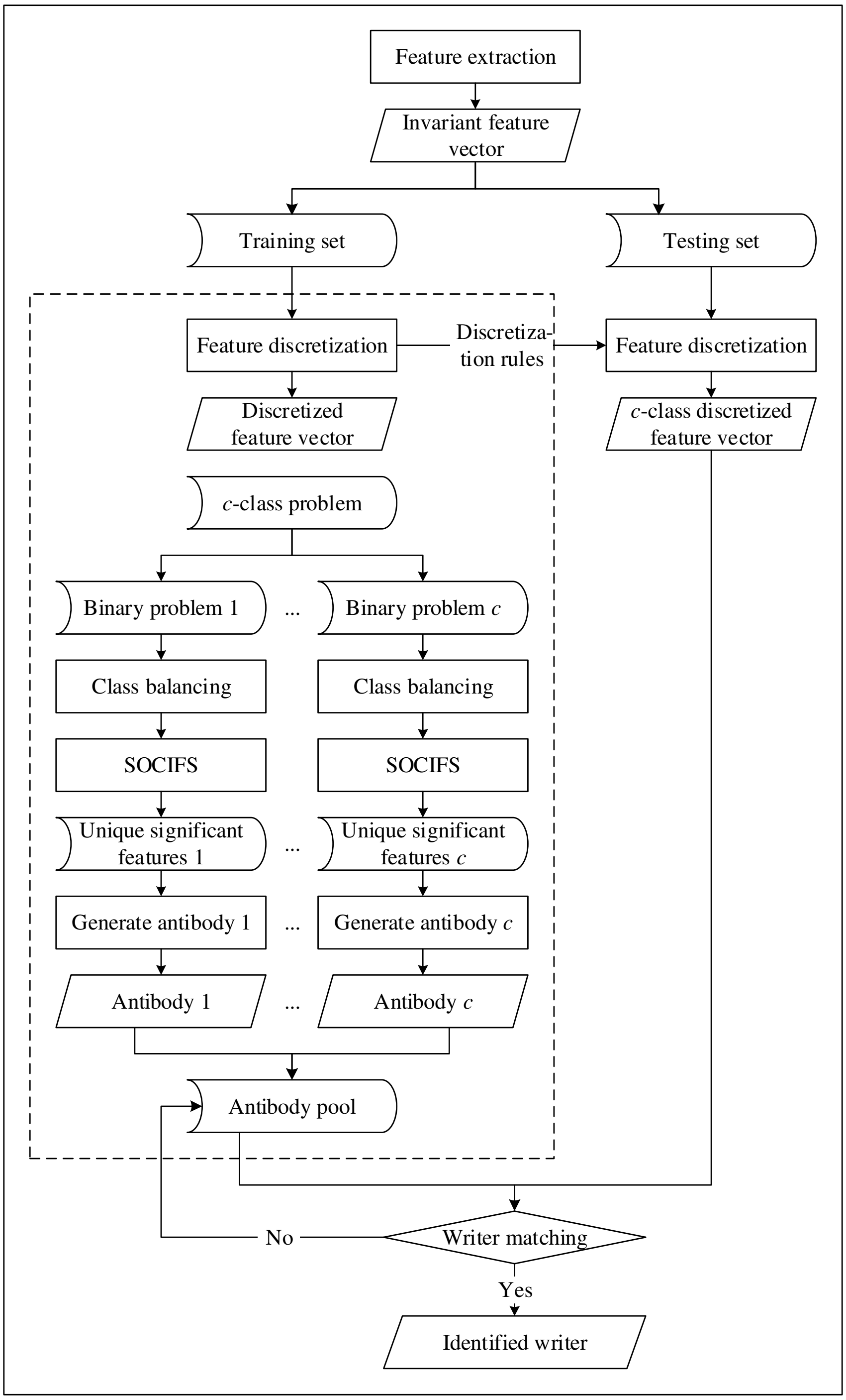


Fig. 7. Cheap Computational Cost Class-Specific Swarm Sequential Selection (C_4S_4)

5 Results and Discussions

The quality of proposed framework must be justified via performance measurements. Dataset used for the performance measurements comes from IAM Handwriting Database [95], which is developed by Research Group on Computer Vision and Artificial Intelligence at Instituts für Informatik und angewandte Mathematik (IAM) in Universität Bern, Switzerland. This database contains forms of handwritten English text. It can be used to train and test handwriting recognition techniques, and to perform writer identification and verification experiments.

Sixty (60) classes are used for research. From these 60 classes, 4400 instances are collected, and are randomly divided into four different datasets to form training and testing dataset in the classification task. The ratio between the number of training and testing dataset is 4:1, which is actually the simple way of describing 5-fold cross-validation. These four datasets are as depicted in Fig. 8.

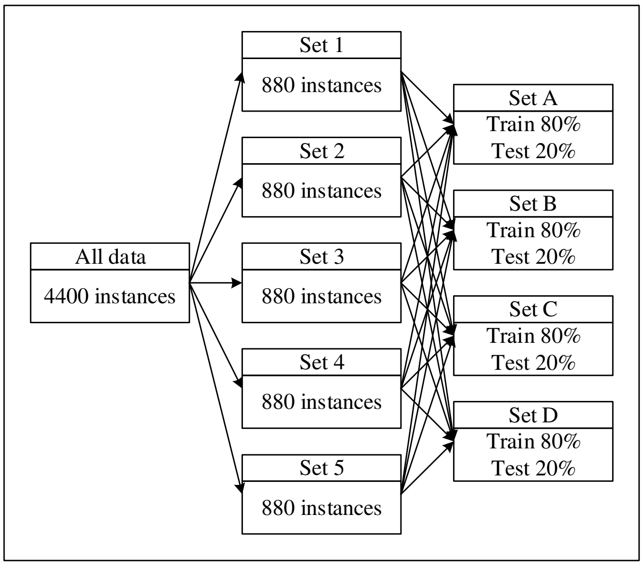


Fig. 8. Data collection procedure

The three commonly used performance measurements for evaluating the performance of feature selection technique are number of selected features, classification accuracy, and processing time. However, considering that the Cheap Computational Cost Class-Specific Swarm Sequential Selection (C_4S_4) will produce different size of feature subset for different class, number of selected features performance measurement will be omitted in this analysis. This analysis will compare the performance of proposed framework to Enhanced Writer Identification Framework (EWIF) and traditional pattern recognition (TPR) framework. Table 3 presents the classification accuracy and processing time results for C_4S_4, EWIF, and TPR in four datasets. The results are also depicted in bar chart format in Fig. 9 and Fig. 10.

Table 3. C4S4, EWIF, and TPR results on classification accuracy and processing time

Criteria	Framework	Set A	Set B	Set C	Set D
Classification accuracy	C_4S_4	99.10%	99.65%	99.09%	99.55%
	EWIF	95.82%	95.65%	95.78%	95.35%
	TPR	45.88%	47.24%	40.14%	39.23%
Processing time	C_4S_4	39.97 sec.	39.05 sec.	39.55 sec.	38.72 sec.
	EWIF	26.97 sec.	24.56 sec.	22.75 sec.	22.91 sec.
	TPR	18.41 sec.	16.29 sec.	16.73 sec.	16.18 sec.

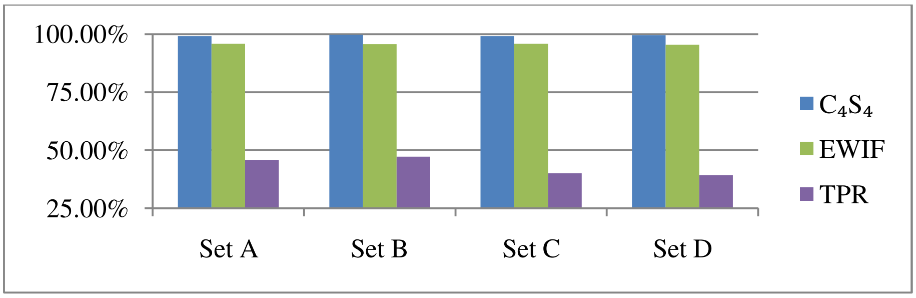


Fig. 9. C4S4, EWIF, and TPR results for classification accuracy

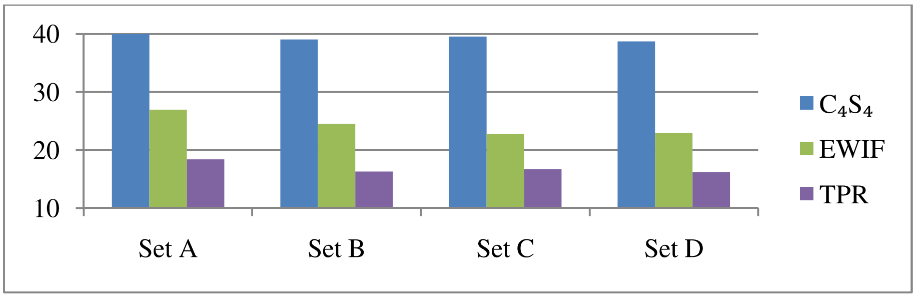


Fig. 10. C4S4, EWIF, and TPR results for processing time

The classification accuracy of proposed framework and feature selection techniques are the primary consideration of this chapter. Based on the results shown in Table 3 and presented graphically in Fig. 9 and Fig. 10, the proposed framework produces the best average of classification accuracy, 99.35%; moreover, the result is significantly exceeding the result of EWIF (95.65%) and TPR (43.12%). The results produced by C_4S_4 shows that the incorporation of feature selection to EWIF is capable to improve its performance. The second measurement of this chapter is processing time of proposed framework and feature selection techniques. Based on the result, it is shown that there is no trade-off between classification accuracy of C_4S_4 and its processing time. The average processing time of C_4S_4 is only approximately 15 seconds longer than EWIF (39.32 to 24.30 seconds) and approximately 23 seconds longer than TPR (39.32 to 16.90 seconds).

6 Conclusions

The purpose of this section is to discuss the summary of this chapter. This chapter is inspired by the fact that every person has unique and significant features that can distinguish oneself to other person, which is always consistent in every handwriting, regardless of words written. These unique individual significant features however, are hidden in the shape and of writing, and thus, key concern in Writer Identification (WI) is in acquiring the features reflecting the author of handwriting using various writing styles.

In this chapter, the word shape is first obtained via feature extraction phase using holistic approach of global representation technique in Moment Function. These extracted features are then selected in the feature selection phase using proposed technique. These selected features are the unique individual significant features which are unique to each person, and used in the classification phase in order to identify the handwritten authorship.

The focus of this chapter is to develop a swarm-based framework which is suited in WI domain, specifically in obtaining the significantly unique features of an individual. The development of the proposed technique and framework has been thoroughly discussed. The proposed framework is unique due to the fact that rather than trying to acquire the features which can differentiate one person to another, the proposed framework instead determine which features are unique to one author. The prior method is commonly used in other domains, where it is important to discriminate one class to another class. However, this is not the case in WI domain. If the prior method is used, the features capable to differentiate one author to another author may not exist, because it is possible for one author possess similar features to another author, although this possibility is rather insignificant. Therefore, the latter method is more suitable, because as mentioned earlier, every individual possess unique and individualistic significant features.

As a conclusion, this chapter has successfully proposed a novel swarm-based framework namely Cheap Computational Cost Class-Specific Swarm Sequential Selection (C_4S_4) which serves as the major contribution of this chapter. While the proposed technique is still not perfect, it still performs better than existing handwritten authorship identification frameworks. The results validate the quality of the proposed technique and framework and open the opportunity for further exploration in WI domain specifically, and other domains generally.

References

1. Amend, K., Ruiz, M.S.: Handwriting Analysis. The Career Press, New Jersey (1980)
2. Huber, R.A., Headrick, A.M.: Handwriting Identification: Facts and Fundamentals. CRC Press, New York (1999)
3. Baranoski, F.L., Oliveira, L.S., Justino, E.J.R.: Writer Identification Based on Forensic Science Approach. In: XXXIII Latin American Conference on Informatics, San Jose, Costa Rica, pp. 25–32 (2007)

4. Muda, A.K.: Authorship Invarianceness for Writer Identification Using Invariant Discretization and Modified Immune Classifier. Universiti Teknologi Malaysia (2009)
5. Li, J., Zheng, R., Chen, H.: From Fingerprint to Writeprint. Communications of the ACM 49(4), 76–82 (2006)
6. Al-Ma'adeed, S., Mohammed, E., Kassis, D.A., Al-Muslih, F.: Writer Identification Using Edge-based Directional Probability Distribution Features for Arabic Words. In: Proceedings of the 2008 IEEE/ACS International Conference on Computer Systems and Applications 2008, pp. 582–590. IEEE Computer Society (2008), 1544403
7. Niels, R., Vuurpijl, L., Schomaker, L.: Automatic Allograph Matching in Forensic Writer Identification. International Journal of Pattern Recognition and Artificial Intelligence 21(1), 61–81 (2007), doi:10.1142/S0218001407005302
8. Pervouchine, V., Leedham, G.: Extraction and Analysis of Forensic Document Examiner Features Used for Writer Identification. Pattern Recognition 40(3), 1004–1013 (2007), doi:10.1016/j.patcog.2006.08.008
9. Srihari, S.N., Huang, C., Srinivasan, H., Shah, V.: Biometric and Forensic Aspects of Digital Document Processing. In: Chaudhuri, B.B. (ed.) Digital Document Processing. Advances in Pattern Recognition, pp. 379–405. Springer, Heidelberg (2007)
10. Tapiador, M., Sigüenza, J.: Writer Identification Method Based on Forensic Knowledge. In: Zhang, D., Jain, A.K. (eds.) ICBA 2004. LNCS, vol. 3072, pp. 555–561. Springer, Heidelberg (2004)
11. Srihari, S.N., Cha, S.-H., Arora, H., Lee, S.: Individuality of Handwriting. Journal of Forensic Sciences, 856–872 (2002)
12. Franke, K., Köppen, M.: A Computer-based System to Support Forensic Studies on Handwritten Documents. International Journal on Document Analysis and Recognition 3(4), 218–231 (2001), doi:10.1007/PL00013565
13. Liu, C.-L., Nakashima, K., Sako, H., Fujisawa, H.: Handwritten Digit Recognition: Benchmarking of the State-of-the-art Techniques. Pattern Recognition 36(10), 2271–2285 (2003)
14. Liu, C.-L., Nakashima, K., Sako, H., Fujisawa, H.: Handwritten Digit Recognition: Investigation of Normalization and Feature Extraction Techniques. Pattern Recognition 37(2), 265–279 (2004)
15. Xu, D.-Y., Shang, Z.-W., Tang, Y.-Y., Fang, B.: Handwriting-based Writer Identification with Complex Wavelet Transform. In: International Conference on Wavelet Analysis and Pattern Recognition, pp. 597–601 (2008)
16. Bensefia, A., Paquet, T., Heutte, L.: A Writer Identification and Verification System. Pattern Recognition Letters 26(13), 2080–2092 (2005)
17. He, Z.-Y., Tang, Y.-Y.: Chinese Handwriting-based Writer Identification by Texture Analysis. In: Proceedings of International Conference on Machine Learning and Cybernetics, pp. 3488–3491 (2004)
18. Yu, K., Wang, Y., Tan, T.: Writer Identification Using Dynamic Features. In: Zhang, D., Jain, A.K. (eds.) ICBA 2004. LNCS, vol. 3072, pp. 512–518. Springer, Heidelberg (2004)
19. Schlapbach, A., Bunke, H.: Off-line Handwriting Identification Using HMM Based Recognizers. In: 17th International Conference on Pattern Recognition, Cambridge, pp. 654–658 (2004)
20. Shen, C., Ruan, X.-G., Mao, T.-L.: Writer Identification Using Gabor Wavelet. In: Proceedings of the 4th World Congress on Intelligent Control and Automation, pp. 2061–2064 (2002)
21. Gupta, S.: Automatic Person Identification and Verification using Online Handwriting. International Institute of Information Technology (2008)

22. Bulacu, M.L.: Statistical Pattern Recognition for Automatic Writer Identification and Verification. University of Groningen (2007)
23. Zhang, B., Srihari, S.N.: Analysis of Handwriting Individuality Using Word Features. In: Proceedings of the Seventh International Conference of Document Analysis and Recognition, pp. 1142–1146 (2003)
24. Marti, U.V., Messerli, R., Bunke, H.: Writer Identification Using Text Line Based Features. In: Proceedings of Sixth International Conference on Document Analysis and Recognition, pp. 101–105 (2001)
25. Srihari, S.N., Cha, S.-H., Arora, H., Lee, S.: Individuality of Handwriting: A Validation Study. In: Sixth IAPR International Conference on Document Analysis and Recognition, pp. 106–109 (2001)
26. Yong, Z., Tieniu, T., Yunhong, W.: Biometric Personal Identification Based on Handwriting. In: Proceedings of 15th International Conference on Pattern Recognition, pp. 797–800 (2000)
27. Zheng, Z., Srihari, R., Srihari, S.N.: A Feature Selection Framework for Text Filtering. In: Proceedings of the Third IEEE International Conference on Data Mining (ICDM 2003), pp. 705–708. IEEE (2003)
28. Srihari, S.H., Cha, S.-H., Lee, S.: Establishing Handwriting Individuality Using Pattern Recognition Techniques. In: Sixth International Conference on Document Analysis and Recognition, pp. 1195–1204 (2001)
29. Schlapbach, A., Kilchherr, V., Bunke, H.: Improving Writer Identification by Means of Feature Selection and Extraction. In: Eight International Conference on Document Analysis and Recognition, pp. 131–135. IEEE (2005)
30. Said, H.E.S., Tan, T., Baker, K.: Personal Identification Based on Handwriting. Pattern Recognition 33, 149–160 (2000)
31. Cha, S.-H., Srihari, S.: Multiple Feature Integration for Writer Verification. In: Proceedings of 7th International Workshop on Frontiers in Handwriting Recognition, pp. 333–342 (2000)
32. Zhang, B., Srihari, S.N., Lee, S.: Individuality of Handwritten Characters. In: Proceedings of 7th International Conference on Document Analysis and Recognition, pp. 1086–1090 (2003)
33. Zois, E.N., Anastassopoulos, V.: Morphological Waveform Coding for Writer Identification. Pattern Recognition 33, 385–398 (2000)
34. Bulacu, M., Schomaker, L., Vuurpijl, L.: Writer Identification Using Edge-based Directional Features. In: Proceedings of 7th International Conference on Document Analysis and Recognition, pp. 937–941 (2003)
35. Schomaker, L., Bulacu, M.: Automatic Writer Identification using Connected-component Contours and Edge-based Features of Uppercase Western Script. IEEE Transactions on Pattern Analysis and Machine Intelligence 26, 787–798 (2004)
36. Bensefia, A., Paquet, T., Heutte, L.: Handwriting Analysis for Writer Verification. In: Proceedings of 9th International Workshop on Frontiers in Handwriting Recognition, pp. 196–201 (2004)
37. Leedham, G., Chachra, S.: Writer Identification Using Innovative Binarised Features of Handwritten Numerals. In: Proceedings of 7th International Conference on Document Analysis and Recognition, pp. 413–417 (2003)
38. Siddiqi, I.: Classification of Handwritten Documents: Writer Recognition. Université Paris Descartes (2009)
39. He, Z., Youb, X., Tang, Y.-Y.: Writer Identification Using Global Wavelet-based Features. Neurocomputing 71(10), 1831–1841 (2008)

40. Yu, L., Liu, H.: Efficient Feature Selection via Analysis of Relevance and Redundancy. Journal of Machine Learning Research, 1205–1224 (2004)
41. Hall, M.A.: Correlation-based Feature Subset Selection for Machine Learning. University of Waikato (1999)
42. Dash, M., Liu, H.: Feature Selection for Classification. Journal of Intelligent Data Analysis, 131–156 (1997)
43. Zhang, P., Bui, T.D., Suen, C.Y.: Feature Dimensionality Reduction for the Verification of Handwritten Numerals. Pattern Analysis Application, 296–307 (2004)
44. Kim, G., Kim, S.: Feature Selection Using Genetic Algorithms for Handwritten Character Recognition. In: Seventh International Workshop on Frontiers in Handwriting Recognition, Amsterdam, pp. 103–112. International Unipen Foundation (2000)
45. Sewell, M.: Feature Selection (2007), `http://machine-learning.martinsewell.com/feature-selection/feature-selection.pdf` (accessed October 25, 2009)
46. Ahmad, A., Dey, L.: A Feature Selection Technique for Classificatory Analysis. Pattern Recognition Letters, 43–56 (2004)
47. Guyon, I., Elisseeff, A.: An Introduction to Variable and Feature Selection. Journal of Machine Learning Research 3, 1157–1182 (2003)
48. Portinale, L., Saitta, L.: Feature Selection: State of the Art. In: Feature Selection, pp. 1–22. Universita del Piemonte Orientale, Alessandria (2002)
49. Kudo, M., Sklansky, J.: Comparison of Algorithms that Select Features for Pattern Classifiers. Journal of Pattern Recognition 33, 25–41 (2000)
50. Yinan, S., Weijun, L., Yuechao, W.: United Moment Invariants for Shape Discrimination. In: International Conference on Robotics, Intelligent Systems and Signal Processing, Changsha, pp. 88–93. IEEE (2003)
51. Hu, M.K.: Visual Pattern Recognition by Moment Invariants. IRE Transactions on Information Theory, 179–187 (1962)
52. Flusser, J., Suk, T., Zitová, B.: Moments and Moment Invariants in Pattern Recognition, vol. 1. John Wiley and Sons, Ltd., West Sussex (2009)
53. Belkasim, S.O., Shridhar, M., Ahmadi, M.: Pattern Recognition with Moment Invariants: A Comparative Study and New Results. Pattern Recognition 24(12), 1117–1138 (1991)
54. Ding, M., Chang, J., Peng, J.: Research on Moment Invariants Algorithm. Journal of Data Acquisition and Processing 7(2), 1–9 (1992)
55. Lv, H., Zhou, J.: Research on Discrete Moment Invariance Algorithm. Journal of Data Acquisition and Processing 1(2), 151–155 (1993)
56. Wang, B., Sun, J., Cai, A.: Relative Moments and Their Applications to Geometric Shape Recognition. Journal of Image and Graphics 6(3), 296–300 (2002)
57. Liu, J., Zhang, T.: Construction and Expansion of Target's Moment Invariants. The Transaction of Photo Electricity Technique 2002, 123–130 (2002)
58. Mukundan, R., Ramakrishnan, K.R.: Moment Functions in Image Analysis Theory and Application. World Scientific Publishing Co. Pte. Ltd., Singapore (1998)
59. Vinciarelli, A.: A Survey on Off-line Cursive Word Recognition. Pattern Recognition, 1433–1446 (2002)
60. Madvanath, S., Govindaraju, V.: The Role of Holistic Paradigms in Handwritten Word Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 149–164 (2001)
61. Agre, G., Peev, S.: On Supervised and Unsupervised Discretization. Cybernetics and Information Technologies 2(2), 43–57 (2002)

62. Nguyen, H.S.: Discretization Problems for Rough Set Methods. In: Polkowski, L., Skowron, A. (eds.) RSCTC 1998. LNCS (LNAI), vol. 1424, pp. 545–552. Springer, Heidelberg (1998)
63. Xin, G., Xiao, Y., You, H.: Discretization of Continuous Interval-valued Attributes in Rough Set Theory and Application. In: International Conference on Machine Learning and Cybernetics, pp. 3682–3686 (2007)
64. Liu, H., Hussain, F., Tan, C.-L., Dash, M.: Discretization: An Enabling Technique. Data Mining and Knowledge Discovery 6, 292–423 (2002)
65. Dougherty, J., Kohavi, R., Sahami, M.: Supervised and Unsupervised Discretization of Continuous Features. In: Proceedings of the Twelfth International Conference on Machine Learning, pp. 194–202 (1995)
66. Kotsiantis, S., Kanellopoulos, D.: Discretization Techniques: A recent survey. GESTS International Transactions on Computer Science and Engineering 32(1), 47–58 (2006)
67. Kohavi, R., John, G.H.: Wrappers for Feature Subset Selection. Artificial Intelligence 97(1-2), 1–43 (1997)
68. Gadat, S., Younes, L.: A Stochastic Algorithm for Feature Selection in Pattern Recognition. Journal of Machine Learning Research, 509–547 (2007)
69. Saeys, Y., Inza, I., Larranaga, P.: A Review of Feature Selection Techniques in Bioinformatics. Journal of Bioinformatics, 2507–2517 (2007)
70. Liu, Y., Wang, G., Chen, H., Dong, H., Zhu, X., Wang, S.: An Improved Particle Swarm Optimization for Feature Selection. Journal of Bionic Engineering 8(2), 191–200 (2011)
71. Unler, A., Murat, A.: A Discrete Particle Swarm Optimization Method for Feature Selection in Binary Classification Problems. European Journal of Operational Research 206, 528–539 (2010)
72. Deriche, M.: Feature Selection Using Ant Colony Optimization. In: 6th International Multi-Conference on Systems, Signals and Devices, pp. 1–4 (2009)
73. Zhuo, L., Zheng, J., Wang, F., Li, X., Ai, B., Qian, J.: A Genetic Algorithm Based Wrapper Feature Selection Method for Classification of Hyperspectral Images Using Support Vector Machine. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 37, 397–402 (2008)
74. Subbotin, S., Oleynik, A.: Modifications of Ant Colony Optimization Method for Feature Selection. In: 9th International Conference - The Experience of Designing and Applications of CAD Systems in Microelectronics, pp. 493–494 (2007)
75. Wang, X., Yang, J., Teng, X., Xia, W., Jensen, R.: Feature Selection based on Rough Sets and Particle Swarm Optimization. Pattern Recognition Letters 28(4), 459–471 (2007)
76. Ververidis, D., Kotropoulos, C.: Fast and Accurate Sequential Floating Forward Feature Selection with the Bayes Classifier Applied to Speech Emotion Recognition. Signal Processing 88(12), 2956–2970 (2008)
77. Xie, J., Xie, W., Wang, C., Gao, X.: A Novel Hybrid Feature Selection Method Based on IFSFFS and SVM for the Diagnosis of Erythemato-Squamous Diseases. In: Workshop on Applications of Pattern Analysis, pp. 142–151 (2010)
78. Chuang, L.-Y., Chang, H.-W., Tu, C.-J., Yang, C.-H.: Improved Binary PSO for Feature Selection using Gene Expression Data. Computational Biology and Chemistry 32(1), 29–38 (2008)
79. Deisy, C., Subbulakshmi, B., Baskar, S., Ramaraj, N.: Efficient Dimensionality Reduction Approaches for Feature Selection. In: International Conference on Computational Intelligence and Multimedia Applications, pp. 121–127 (2007)

80. Pizzi, N.J., Pedtycz, W.: Classification of Magnetic Resonance Spectra using Parallel Randomized Feature Selection. In: Proceedings of 2004 IEEE International Joint Conference on Neural Networks, pp. 2455–2459 (2004)
81. Melab, N., Cahon, S., Talbi, E.-G.: Parallel GA-based Wrapper Feature Selection for Spectroscopic Data Mining. In: Proceedings of International Parallel and Distributed Processing Symposium, pp. 201–208 (2002)
82. Qian, X.-J., Xu, J.-B.: Optimization and Implementation of Sorting Algorithm Based on Multi-core and Multi-thread. In: IEEE 3rd International Conference on Communication Software and Networks, pp. 29–32 (2011)
83. Cruz, C., Pelta, D.A., Royo, A.S., Verdegay, J.L.: Soft Computing and Cooperative Strategies for Optimization. In: IEEE Mid-Summer Workshop on Soft Computing in Industrial Applications 2005, pp. 75–78 (2005)
84. Pudil, P., Novovicova, J., Kittler, J.: Floating Search Methods in Feature Selection. Pattern Recognition Letters 15, 1119–1125 (1994)
85. Whitney, A.W.: A Direct Method of Nonparametric Measurement Selection. IEEE Transaction in Computational, 1100–1103 (1971)
86. Eberhart, R.C., Kennedy, J.: A New Optimizer using Particle Swarm Theory. In: Proceedings of 6th International Symposium on Micro Machine and Human Science, pp. 39–43 (1995)
87. Kennedy, J., Eberhart, R.C.: Particle Swarm Optimization. In: Proceedings of IEEE International Conference on Neural Networks 1995, pp. 1942–1948 (1995)
88. Abdl, K.M., Mohd Hashim, S.Z.: Swarm-Based Feature Selection for Handwriting Identification. Journal of Computer Science 6(1), 80–86 (2010)
89. Pratama, S.F., Muda, A.K., Choo, Y.-H., Muda, N.A.: A Comparative Study of Feature Selection Methods for Authorship Invarianceness in Writer Identification. International Journal of Computer Information Systems and Industrial Management Applications 4, 467–476 (2012)
90. Pratama, S.F., Muda, A.K., Choo, Y.-H., Muda, N.A.: PSO and Computationally Inexpensive Sequential Forward Floating Selection in Acquiring Significant Features for Handwritten Authorship. In: 11th International Conference on Hybrid Intelligent Systems, Melaka, Malaysia, pp. 358–363 (2011)
91. Pratama, S.F., Muda, A.K., Choo, Y.-H., Muda, N.A.: Computationally Inexpensive Sequential Forward Floating Selection for Acquiring Significant Features for Authorship Invarianceness in Writer Identification. International Journal of New Computer Architectures and Their Applications 1(3), 581–598 (2011)
92. Pratama, S.F., Muda, A.K., Choo, Y.-H., Muda, N.A.: SOCIFS Feature Selection Framework for Handwritten Authorship. International Journal of Hybrid Intelligent Systems 10(2), 83–91 (2013), doi:10.3233/HIS-130167
93. Muda, A.K., Shamsuddin, S.M., Darus, M.: Invariants Discretization for Individuality Representation in Handwritten Authorship. In: Srihari, S.N., Franke, K. (eds.) IWCF 2008. LNCS, vol. 5158, pp. 218–228. Springer, Heidelberg (2008)
94. Pineda-Bautista, B.B., Carrasco-Ochoa, J.A., Martinez-Trinidad, J.F.: General Framework for Class-Specific Feature Selection. Expert Systems with Applications 38, 10018–10024 (2011)
95. Marti, U., Bunke, H.: The IAM-database: an English Sentence Database for Off-line Handwriting Recognition. International Journal on Document Analysis and Recognition 5, 39–46 (2002)