*Journal of Information Hiding and Multimedia Signal Processing* ©2024 ISSN 2073-4239
**Ubiquitous International** Volume **15**, Number **4**, December **2024**

# Handwriting Based Writer Identification Using a Fragment Encoding System

Abdelillah Semma, Said Lazrak

Faculty of Sciences
Ibn Tofail University
Campus Universitaire, BP 133, Kenitra, Morocco
semma_abdelillah@yahoo.fr;said.lazrak@uit.ac.ma

Yaâcoub Hannad

Faculty of Educational Sciences
Mohammed V University
Madinat Al Irfane, B.P. 6211, Rabat, Morocco
y.hannad@gmail.com

Received September 2024, revised November 2024, accepted November 2024.

Abstract. *The identification of authorship in handwritten documents presents a challenging task in the field of pattern recognition. The most commonly used codebook-based approaches for capturing writing style employ randomly extracted fragments. These methods may not fully represent the nuances of the writing style, and require significant computational resources and time due to the numerous comparison operations between writing fragments in test and reference documents. This paper proposes an efficient offline writer recognition system that addresses the aforementioned limitations by using VLAD encoding of fragments extracted from multiple keypoints, including FAST, SIFT, SURF, and Harris. Recently introduced encoding techniques, such as VLAD, have shown improved classification results by aggregating local descriptor residuals based on the nearest cluster. Our study demonstrates that encoding redundant patterns within small fragments can outperform established keypoint descriptors like SIFT and SURF. The proposed approach involves extracting small fragments centered around various keypoints, followed by applying the VLAD encoding method on flattened fragments to generate a global descriptor for each image. We conducted experiments on five public datasets (BFL, CVL, IAM, QUWI-EN, and QUWI-AR) and obtained promising results, with identification rates of 100%, 100%, 97%, 98.3%, and 96.9%, respectively.*
**Keywords:** Writer Identification, keypoint, VLAD, Encoding Fragment, Codebook.

1. **Introduction.** Handwriting has been one of the ancient means of communication, alongside audiovisual tools. Learning to write from an early age allows individuals to develop their distinct writing style. However, variations in writing instruments, writing surfaces, and an individual's physical and mental state significantly affect writing styles. Furthermore, even under identical circumstances and with the same tools, producing two identical writings is impossible, contributing to the complexity of writer identification tasks.

Despite this complexity, scientific research has explored the potential correlation between a writer's personality and their handwriting [1], as well as the relationship between the handwritten text and the writer's gender or age [2]. Some studies have even helped predict the date of writing for historical manuscripts [3].

Advancements in image processing and pattern recognition have led to the development of automatic systems for analyzing and classifying images and handwritten documents. These systems seek to translate the knowledge base of forensic experts into digital tools and automatic systems, used to identify the true author of a handwritten document. Such systems can assist experts by proposing a list of likely candidates from a large dataset.

Among the various known approaches for identifying writers of manuscript documents, codebook-based techniques have garnered considerable attention from researchers. These methods rely on the extraction of a vocabulary of models, known as codebooks, which allow for the effective representation of the characteristics studied. This is because, during the writing process, individuals tend to use similar strokes and basic shapes generated by the same hand gestures, resulting in the generation of redundant forms of writing with varying frequencies. These redundant forms, or "codebooks", represent invariant writing traits of each writer, and have shown promising results in the field of writer identification.

However, despite the successes achieved, codebook-based writer identification systems suffer from several limitations. Firstly, these systems rely on randomly extracted fragments that are grouped into separate classes according to their similarity, with a single fragment chosen as the representative of each class. This approach may result in the loss of important information contained in other fragments of the same class. Secondly, the classification process is time-consuming due to the high number of comparison operations required between the codebooks of test and reference documents. Additionally, the random extraction of fragments may not accurately represent the writer's writing style.

To address these issues, we propose the use of keypoints or corners as locations for fragment extraction. These keypoints represent locations where the writing suddenly changes direction, and the forms of writing around these points are more likely to be more discriminatory in characterizing the author than random fragments. Our second contribution involves using the Vector of Locally Aggregated Descriptors (VLAD) encoding technique to aggregate the various local descriptors to obtain a single global descriptor per handwritten document. This approach enables the consideration of all the extracted fragments and considerably reduces the execution time while ensuring higher classification rates.

The highlights of our study include:

- Proposing an identification system based on textural features extracted from square patches centered around several keypoints.
- Conducting a comparative study between methods based on the encoding of keypoint descriptors and those based on the encoding of small fragments.
- Demonstrating the efficiency of selecting keypoints as extraction locations by comparing them with randomly extracted fragments.
- Performing experiments on five public datasets written in four languages.

The remainder of this paper is organized as follows: In Section 2, we review relevant literature on writer identification. In Section 3, we describe our adopted approach and the methods employed. In Section 4, we present the test results, along with analysis and discussion. Finally, we conclude the paper with a brief summary and discuss future perspectives.

2. **Related Work.** In an important study by [4], which used handwritten documents of 1500 writers, the individuality of handwriting was demonstrated, establishing its potential as an effective biometric tool. Subsequently, a number of writer recognition systems were developed based on two main groups of features: textural and structural. Structural features, calculated at global or local levels, aim to capture the structural characteristics of writing such as average intra-word and inter-word distances, line height, and writing

inclination. While a number of writer identification methods based on structural features [5, 6] have recorded high classification results, they require a long execution time due to the complexity of segmentation and feature extraction steps.

Texture analysis-based techniques rely on retrieving a set of characteristics from specific regions [7] or the entire image [8, 9]. Texture-based techniques are known for their short execution time, mainly due to the fast calculation of textural characteristics, unlike structural features. [8] proposed an interesting system that achieved high identification rates by using two local texture features on normalized image regions, namely the Local Phase Quantization (LPQ) [10] and Local Binary Patterns (LBP) [11] descriptors. These local texture features have been widely used in various texture classification systems [8, 10, 11, 12].

In addition to the LBP and LPQ descriptors, other types of descriptors have been proposed, such as Histograms of Oriented Gradients (HOG), which have been extensively used in the field of face recognition [13, 14]. HOG descriptors, which are based on the calculation of histograms from the angles formed by the gradients of vertical and horizontal pixels around a given point, were used by [15] to extract local descriptors from image fragments. The same authors proposed another system [16] by combining HOG descriptors with GLRL White and GLRL Black. Similarly, [9] demonstrated the superiority of local Gray Level Run Length (GLRL) descriptors over Gray Level Co-occurrence Matrices (GLCM) and concluded that GLRL histograms contain more discriminative information. The good performance of textural descriptors has led scientific researchers to test other types of descriptors, such as Oriented Basic Image Features (OBIF), Local Ternary Patterns (LTP) [16], Run length [17], Contour-hinge [5], Contour-direction [5], edge-direction [17] and edge-hinge [17].

Several studies have utilized the discriminative power of redundant patterns to improve writer identification systems. For instance, [18] employed a technique that divided each handwritten image into small fragments of size 13x13 and grouped morphologically similar fragments within the same class. In an extension of this study, [6] exploited two visual elements of writings, orientation and curvature, in addition to the codebook generated from small patches. Similarly, [19] proposed a technique that extracted invariant forms from each handwritten document by dividing Tunisian and Algerian city words into small patches of size 19x19, followed by a comparison of similarity between the invariant forms of the test documents and those of training. In contrast to these studies, [15] utilized a technique based on the extraction of textural descriptors (LBP, LTP, and LPQ) using larger fragments (100x100). In another interesting study, [20] proposed a writer recognition system based on constructing a codebook from small fragments extracted around Harris keypoints. The experimental evaluation on the CVL and BFL datasets achieved an identification rate of 88.5% and 97.3%, respectively, highlighting the effectiveness of exploiting redundant patterns in improving writer identification systems.

Recently, in the last few years, increasing and particular attention has been paid by scientific researchers to deep learning-based approaches [3, 21, 22, 23, 24, 25] thanks to their ability in tasks of classification and identification of complex images. Indeed, deep networks have the capacity to learn discriminating information directly from the data and without going through traditional functionalities based on the calculation of textural or structural descriptors of handwritten images. In 2015, Fiel and sablating [22] designed the first writer identification system using a CaffeNet model with eight layers. In [3], the author proposed an unsupervised deep learning system trained on 32x32 patches centered around SIFT keypoints with labels relating to clustred SIFT descriptors. The same author has proposed another system [23] in which he uses the VLAD encoding of the penultimate layer activations of the ResNet-34 residual neural network to identify the writers of the KHATT and ICDAR13 datasets. In [26], an other end-to-end deep-learning system for text-independent writer recognition was proposed using VLAD encoding method. A global-context residual recurrent neural network (GR-RNN) was employed in [24] using binary images, gray images and handwriting outlines to demonstrate that all textural information contained in a handwritten image constitutes an element important in writer identification systems. In [27], the author proposed a deep neural network (FragNet) trained on words or text blocks of handwritten images. In [28], a study carried out on the bilingual QUWI dataset with the activations of the different layers of the AlexNet model, allowed to achieve scores of 92.78% and 92.20% on the English and Arabic subset respectively.

Inspired by the performance of the VLAD encoding method and the keypoints used in our recent study [29] and by the potential of fragment-based approaches [6, 16, 18, 19, 20, 30], we propose a system based on the combination of these three points by extracting fragments centered around several keypoints which are then encoded via the VLAD method. In the following section we present the details of the proposed methodology.

3. **METHODOLOGY.** The proposed methodology consists of three main steps: feature extraction, feature encoding, and classification. An overview of the approach is shown in Figure 1.

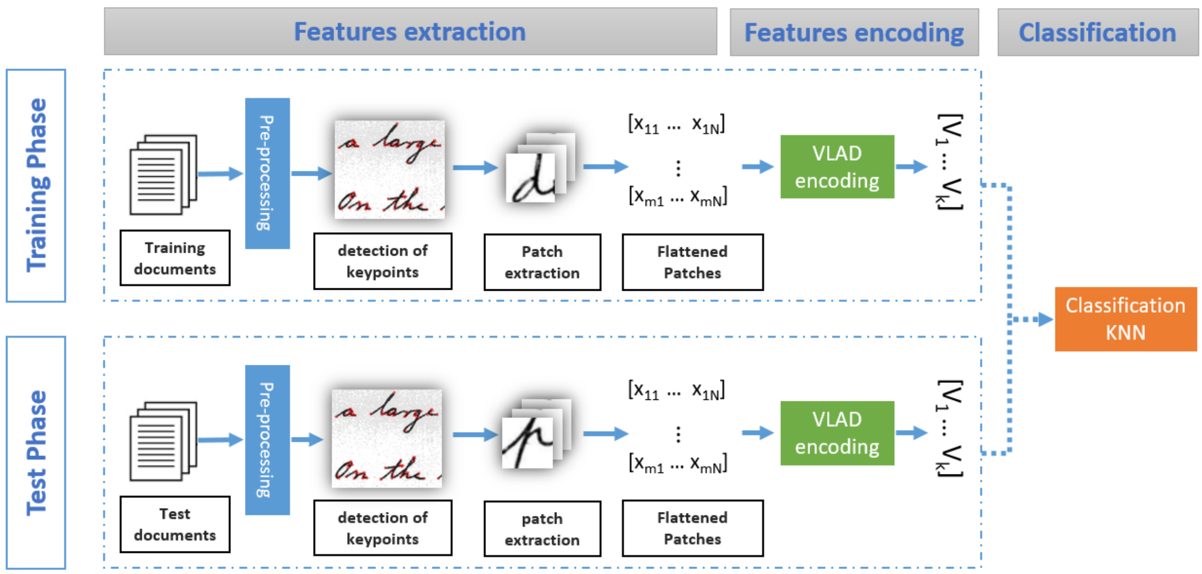


FIGURE 1. Overview of the proposed methodology for writer identification.

For each handwritten image, local descriptors are calculated based on square fragments centered around various types of keypoints. During the learning phase, a dictionary is built from local descriptors (i.e., flattened fragments) of handwritten documents. The VLAD encoding step assigns a global descriptor to each handwritten document based on the dictionary and local descriptors. This global descriptor is then used in the classification step.

3.1. **Features Extraction.** Keypoint-based image classification systems have demonstrated high efficacy in tasks such as face recognition and writer identification. Among these systems, Scale-Invariant Feature Transform (SIFT) is one of the most widely recognized and successful methods. The SIFT method, as described in Lowe's publication titled "Object Recognition from Local Scale-Invariant Features" [31], identifies key points by analyzing the derivatives of the scale space using a method called Gaussian Difference (DoG). Following detection of these points, descriptors are generated utilizing information pertaining to the amplitude and orientation of the gradient.

Despite its high performance, the detection process of Scale-Invariant Feature Transform (SIFT) is known to be comparatively slow. In order to address this issue and achieve faster detection of keypoints, Bay, H., Tuytelaars, T., and Van Gool, L. proposed a novel technique named Speeded Up Robust Features (SURF) in their 2006 publication [32]. This technique is essentially an accelerated version of SIFT and was designed to improve the speed of feature detection.

Prior to the development of Scale-Invariant Feature Transform (SIFT) and Speeded Up Robust Features (SURF), an alternative keypoint detection method was introduced in 1988. Chris Harris and Mike Stephens built upon the concept introduced by Moravec to develop the Harris corner detection algorithm, which is regarded as one of the most notable intensity-based corner detectors. The algorithm calculates the sum of squared differences (SSD) between the original image and its shifted version in any direction at the local level. This technique was presented in their paper titled "A Combined Corner and Edge Detector" [33].

In 2006, Rosten et al. introduced a highly efficient keypoint detection method called Features from Accelerated Segment Test (FAST), which was further improved upon in 2010. This algorithm operates by computing a score that reflects the difference in intensity between the central point and the 16 points located on a circle with a radius of 3 around it. This score-based approach allows for the rapid identification of keypoints, as described in the publication titled "Machine learning for high-speed corner detection" [34].

Once key points are detected, small square patches are extracted around these key points. These patches must be large enough to capture sufficient information about the author's style, but small enough to contain only redundant information or patterns (such as writing strokes and gestures). These flattened patches are then passed through an encoding step, as detailed in the next section.

Figure 2 shows some samples of fragments extracted around the Harris Corner Detector, which are then flattened (as shown in Figure 3) before following the encoding step.

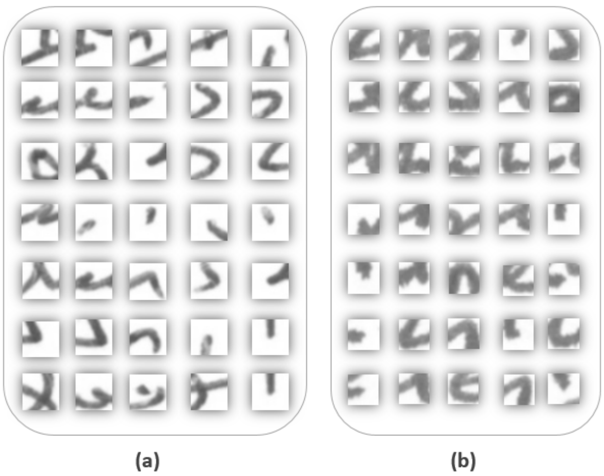


FIGURE 2. Sample Harris Keypoint fragments from (a) QUWI-AR and (b) QUWI-EN.

3.2. **Features Encoding.** The local features extracted from the flattened fragments are encoded to generate a global descriptor for each image. The Vector of Locally Aggregated

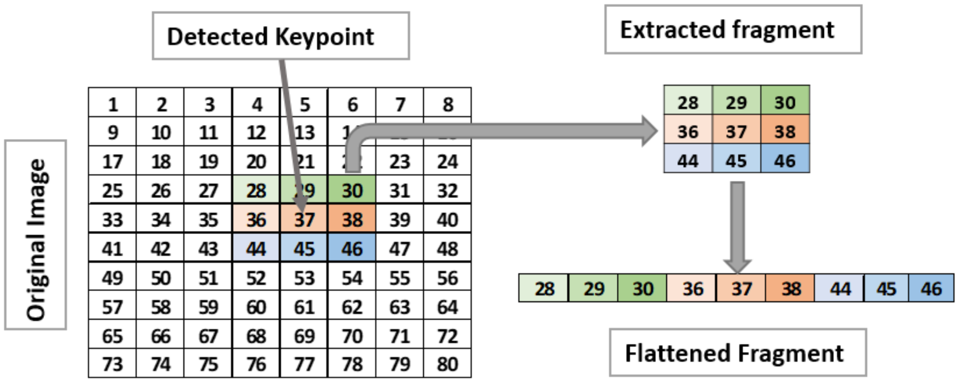


FIGURE 3. Local features extraction process.

Descriptors (VLAD)[35] is a non-probabilistic version of the Fisher kernel[36] that codes the first-order statistics by computing the residuals of the local features with respect to their nearest cluster center. VLAD has demonstrated its effectiveness in several fields, such as face recognition [37, 38] and writer identification [3, 23, 26] and is considered a standard coding algorithm.

To construct the VLAD vector, the dictionary of $k$ clusters ($c_1$, $c_2$, ... $c_k$) is generated using the K-means algorithm. Then, for each cluster center $c_i$, the embedding vector $v_i$ is computed by aggregating the residuals of the local descriptors that are closest to $c_i$:

$$v_i = \sum_{NN(x_j)=c_i} (x_j - c_i) \tag{1}$$

The concatenation of all the $v_i$ represents the global descriptor $\hat{\Psi}$.

$$\hat{\Psi} = (v_1, v_2, ...v_k) \tag{2}$$

Normalization has been shown to enhance the discriminative ability of VLAD encoding. Among the various normalization techniques, power normalization has been widely employed and has demonstrated its effectiveness in several research studies [3, 23, 26]:

$$\Psi = sign(\hat{\Psi}) * |\hat{\Psi}|^p \tag{3}$$

We choose the value 0.3 for the power factor $p$. Finally, we apply the $L2$ normalization for each vector.

$$\Phi = \frac{\Psi}{||\Psi||_2} \tag{4}$$

3.3. **Classification.** The classification step is carried out using the nearest neighbor classification method (KNN), which is based on finding the minimum distance between the test vector and the closest candidates in the training database. Nearest neighbor search (NN) is a popular classification method in computer vision applications, where the task is to find the k closest points to a query point q or to find all points whose distance from q is less than a radius r. To process these queries, a data structure based on space partitioning can be used, such as the Ball tree algorithm [39].

The Ball tree algorithm is a binary decision tree, where each node represents a ball or hyper-sphere containing child balls. To construct the ball tree, the data is first divided into two hyper-spheres, each containing a group of points based on their distances from the center of gravity of the two groups. Each of the two hyper-spheres is then further divided into two child hyper-spheres, and this process is repeated recursively.

4. **Experiments & Results.** In this section, we present the results of the tests conducted to validate the proposed approach, along with a detailed discussion. We begin by introducing the various standard metrics used in the evaluation, followed by a description of the databases employed in the study. Subsequently, we present a comparison between the approach based on keypoints and the approach based on the random extraction of small fragments. Next, we discuss the advantages of encoding patches over encoding descriptors for different keypoints. Finally, we compare the performance of the proposed approach with other techniques.

To evaluate the classification performance of the proposed approach, we use several standard metrics, including Top-N, Soft-N, and Hard-N.

- $Top - N$ : corresponds to the scenario where a training document similar to the query document is classified at rank $N$.
- $Soft - N$ : corresponds to the case where at least one document in the training database similar to the query document is ranked at $N$ or lower.
- $Hard - N$ : corresponds to the case where all the documents in the training database that are ranked at $N$ or lower belong to the correct writer.

4.1. **Datasets.** The five databases used in this study cover a variety of languages and writing styles, which allows for a comprehensive evaluation of the proposed approach. These datasets include two English datasets, namely IAM [40] and QUWI-EN [41], one Arabic dataset, QUWI-AR [41], one Portuguese dataset, BFL [42], and one hybrid language dataset, CVL [43]. Each dataset is described in detail below.

The **IAM dataset** [40] consists of handwritten documents from 657 English writers. Among them, 356 writers produced only one page while the other 301 wrote two or more pages. Consistent with the approach utilized by [5, 44], we have opted to retain the first two writing pages submitted by writers who have contributed more than two pages to our dataset. Alternatively, for writers who have submitted only one page, we have divided the page into two parts. These modifications have resulted in a dataset consisting of handwriting samples contributed by 657 writers, each represented by two samples. Specifically, one writing page of each writer is utilized for training purposes while the second page is reserved for testing.

**The CVL database** [43] includes handwritten documents in English and German from 311 authors. Twenty-seven writers produced seven text documents while the remaining 284 writers produced only five documents. Following the experimental methodology utilized in previous studies such as [44, 45, 46], we have selected the initial four English handwriting samples per writer. These samples have been allocated such that three of them are used in the training set and the fourth is employed in the test set.

**The brazilian BFL database** [42] includes 945 handwritten documents in Portuguese from 315 writers. In accordance with the experimental methodology utilized in prior research, such as [47], we have utilized the entire BFL dataset consisting of three handwriting samples per writer. Specifically, two of the samples have been designated for training purposes while the remaining sample has been allocated for testing.

**The QUWI dataset** [41] is a bilingual database that contains handwritten documents from 1017 writers, where each writer wrote four samples, two in English and two in Arabic. The English version includes only a sample of writings from 975 writers, while the Arabic version includes all 1017 writers. The publicly accessible version of the dataset includes pages divided into paragraphs, generating six Arabic pages for each writer. Consistent with the experimental approach employed in earlier studies such as [26], we have designated one page from each set of handwriting samples for use in the testing stage while the remaining pages have been reserved for training purposes.

Writing samples from QUWI, IAM, BFL, and CVL datasets are presented in Figure 4.

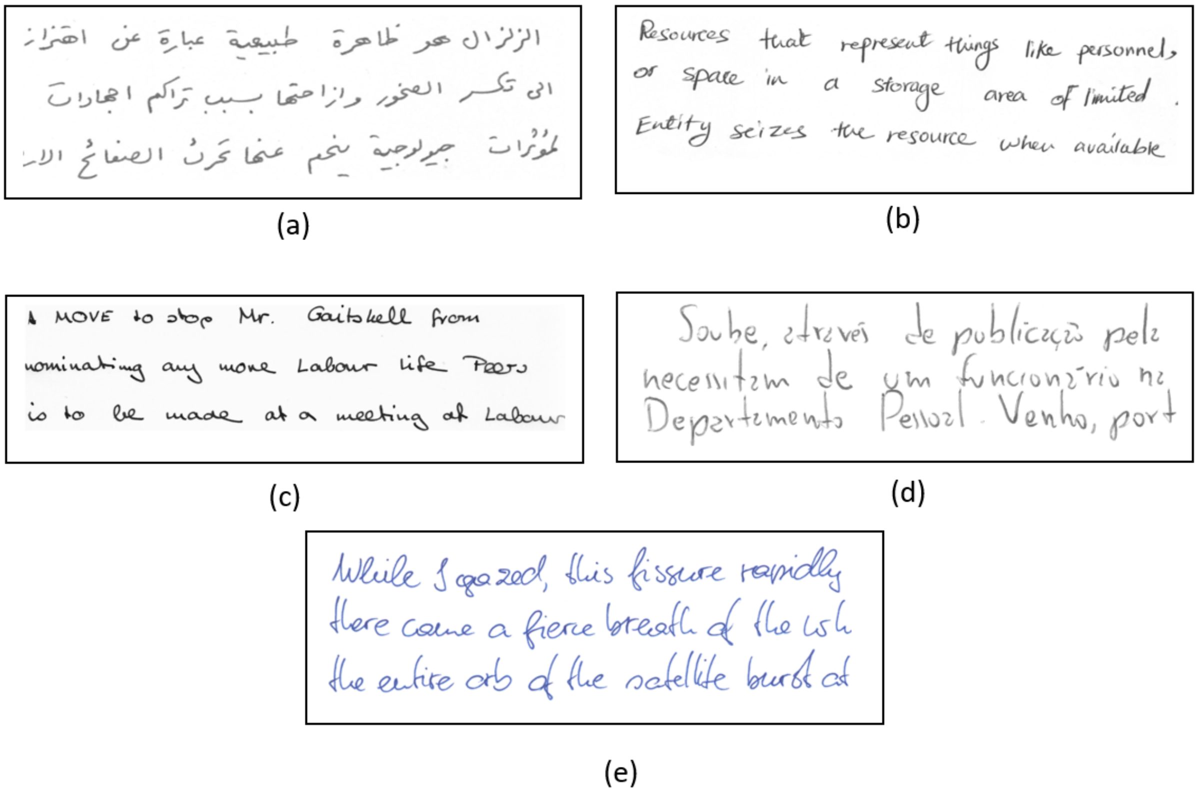


Figure 4. Examples of Handwritten Documents from (a) QUWI-Ar, (b) QUWI-En, (c) IAM, (d) BFL and (e) CVL

4.2. **Sensitivity to the number of writing fragments.** Figure 5 presents the results of the proposed system using the QUWI (AR) database with several numbers of 39x39 fragments and several extraction methods.

The study results confirm the positive correlation between the amount of training data and the system's performance. Irrespective of the method used, the increase in the number of learning fragments significantly enhances the performance of the proposed approach.

Furthermore, the results indicate that there is a substantial gap in the identification rates obtained using the random extraction method and those based on keypoints for small numbers of fragments (100 and 500). Specifically, the use of 100 randomly extracted fragments resulted in a Top-1 identification rate of 7.1%, while the use of fragments centralized around the Harris Corner detector resulted in a rate of 46.2%. The rates were even higher than 30% for the case of fragments centralized around the other key points (FAST, SIFT and SURF). These results highlight the superior discriminative power of the fragments extracted around keypoints compared to those extracted randomly.

4.3. **Sensitivity to window size.** Our proposed approach for writer identification is based on encoding handwritten text fragments using VLAD, and the system's performance is logically sensitive to the size of the selected fragments. Table 1 presents the Top-1 identification rates calculated for each dataset using various fragment sizes. The

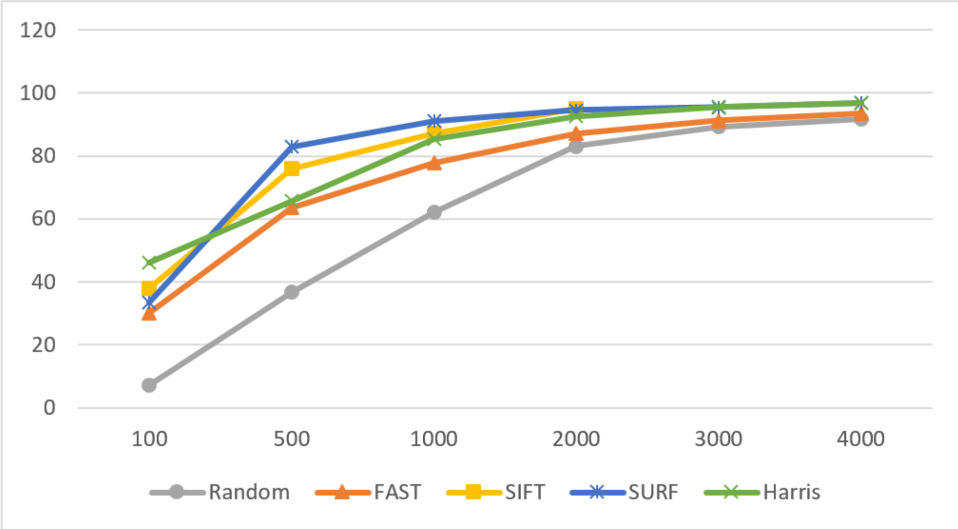


FIGURE 5. Impact of the number fragments on QUWI (AR) database using VLAD encoding **(k = 64)** with multiple extraction methods

TABLE 1. *Top-1 identification rates as a function of window size using fragments centered around SURF key points (Hessian threshold=3000) with the VLAD encoding method* ***(k = 64)***

| Fragment Size | CVL | BFL | IAM | QUWI (AR) | QUWI (EN) |
|---|---|---|---|---|---|
| 45x45 | 99.4 | 99.4 | 92.8 | 93 | 96.5 |
| 42x42 | 99.4 | 99.4 | 93.8 | 92.6 | 97.2 |
| 39x39 | 99.7 | 99.4 | 93.4 | **94.3** | 96.5 |
| 36x36 | 99.7 | 99.4 | **94.2** | 94.1 | 97.1 |
| 33x33 | **100** | **99.4** | 94.2 | 93 | **97.3** |
| 30x30 | 99.4 | 99.4 | 93.4 | 93.9 | 97 |
| 25x25 | 99.4 | 99.4 | 91.9 | 90.9 | 96.8 |
| 20x20 | 99.4 | 98.4 | 91.5 | 90.4 | 96.7 |
| 15x15 | 99 | 98.4 | 90.7 | 87.3 | 95.7 |
| 10x10 | 98.1 | 97.8 | 86.7 | 73.1 | 60.5 |
| 5x5 | 89.7 | 95.2 | 70.1 | 40.3 | 34.5 |

classification rates are generally low for small window sizes ($5x5$ and $10x10$), whereas the system is more stable for medium or large size fragments.

The optimal fragment size varies across datasets. For CVL and QUWI (EN), the best $Top-1$ classification rate is achieved using fragment sizes of $33x33$, while for QUWI (AR) and IAM, the best performance is realized using fragment sizes of $39x39$ and $36x36$, respectively. For BFL, the best identification results are obtained using fragment sizes ranging from $25x25$ to $45x45$.

Compared to other writer identification systems that use small fragments (such as [6],[18], and[19]) or large fragments (such as [16] and [30]), our proposed system employs medium-sized fragments that aim to represent handwriting gestures and strokes likely to be common across multiple writings by the same person.

In the following sections of this paper, we choose the following fragment sizes for each dataset: $33x33$ for BFL, CVL, and QUWI (EN), $36x36$ for IAM, and $39x39$ for QUWI (AR).

4.4. **keypoints features vs Keypoints fragments vs random fragments.** Encoding methods such as VLAD, Fisher Vector, and Bag of Words have been widely used in the encoding of keypoint descriptors [35, 48, 49, 50]. In the proposed approach, we utilize the VLAD encoding of script fragments representing local features of each handwritten document.

To evaluate the effectiveness of our approach in a mono-script environment, we present a performance comparison of the results obtained using encoding of script fragments with those using original descriptors of BRIEF, SIFT, SURF, and ORB in Tables 2 and 3. The Top-1 metric indicates that the performance of fragment encoding significantly outperforms that of descriptor encoding on almost all datasets. Particularly, in the QUWI (AR) dataset, the difference in performance between the two approaches is very clear because no keypoint descriptor except SURF reached 90% as opposed to fragment-based methods (SIFT-Patch, Harris-Patch, FAST-Patch, SURF-Patch, and Random-Patch). Moreover, the difference between the performance of fragments centralized around SIFT keypoints and those of classic descriptors of the same SIFT keypoints is about 10 points.

Table 2. *Classification performance on QUWI (AR) and QUWI (EN) using VLAD encoding of (patches centered around keypoints vs random patches vs keypoint descriptors)*

| | QUWI (AR) | | | QUWI (EN) | | |
|---|---|---|---|---|---|---|
| Extraction method | Top-1 | Hard-2 | Soft-5 | Top-1 | Hard-2 | Soft-5 |
| SIFT-Patch | 95.2 | 81.1 | 98.1 | 97 | 85.3 | **99.3** |
| SURF-Patch | 96.8 | **86.4** | 98.7 | **98.3** | **88.4** | 99.2 |
| Harris-Patch | **96.9** | 84.7 | **99.2** | 97.3 | 84.6 | 99.2 |
| FAST-Patch | 93.5 | 80.6 | 97.9 | 97.1 | 82.4 | 98.7 |
| Random-Patch | 91.9 | 75.8 | 97.4 | 95.4 | 82.1 | 97.9 |
| BRIEF-Desc | 88.6 | 67.6 | 95 | 93.9 | 75.6 | 98.6 |
| SIFT-Desc | 85.4 | 64.8 | 96.3 | 92.9 | 72.6 | 98.2 |
| SURF-Desc | 94.2 | 80.5 | 98.7 | 96.8 | 81.8 | 98.9 |
| ORB-Desc | 87.7 | 66.7 | 96 | 91.9 | 63.9 | 97.8 |

Table 3. *Classification performance on IAM, CVL and BFL using VLAD encoding of (patches centered around keypoints vs random patches vs keypoint descriptors)*

| | IAM | | | CVL | | | BFL | | |
|---|---|---|---|---|---|---|---|---|---|
| Method | Top-1 | Hard-2 | Soft-5 | Top-1 | Hard-2 | Soft-5 | Top-1 | Hard-2 | Soft-5 |
| SIFT-Patch | 96.5 | - | 97.9 | **100** | **99.4** | **100** | **100** | 95.6 | **100** |
| SURF-Patch | 94.4 | - | 97.6 | **100** | **99.4** | **100** | 99.7 | 97.1 | **100** |
| Harris-Patch | 94.2 | - | 97.9 | **100** | **99.4** | **100** | 99.4 | 97.1 | 99.7 |
| FAST-Patch | **97** | - | 97.9 | **100** | 99 | **100** | **100** | 96.8 | **100** |
| Random-Patch | 95.7 | - | **98** | **100** | 99 | **100** | 99 | 96.2 | 99.4 |
| BRIEF-Desc | 89 | - | 97.1 | **100** | 97.4 | **100** | 99.7 | 97.1 | 99.7 |
| SIFT-Desc | 93.8 | - | 97 | **100** | 98.4 | **100** | 99.4 | 97.1 | **100** |
| SURF-Desc | 93.4 | - | 97.4 | **100** | 98.1 | **100** | 99 | 95.2 | 99.7 |
| ORB-Desc | 95 | - | 97.9 | 99.7 | 98.1 | **100** | 98.7 | 94.6 | 99.7 |

Furthermore, we observe that the performance of keypoints varies from one dataset to another. FAST keypoints perform well in the three datasets IAM, CVL, and BFL, while SURF performs well on CVL and on the English version of the QUWI dataset. Similarly, Harris keypoints yield higher Top-1 identification rates in both QUWI(AR) and CVL datasets. The excellent performance of methods based on patches suggests that their descriptors may serve as a potential replacement for classic descriptors of keypoints, despite their larger dimensions.

In addition to the comparison with methods based on descriptors, we also compare the efficiency of our method based on encoding of fragments extracted around keypoints with that of randomly extracted fragments. Tables 2 and 3 demonstrate the identification rates reported by the two approaches, where we can see that the method based on keypoint fragments significantly outperforms the random extraction method. However, we observe that encoding of randomly extracted fragments also yields good results on the CVL and BFL datasets characterized by their medium number of classes, unlike the QUWI (AR) and QUWI (EN) datasets where the random extraction method yields lower identification rates. These results demonstrate the usefulness of using keypoints as extraction points, as these points of interest define the specific and interesting information contained in each image.

4.5. **Performance Comparison with State-of-the-art.** The recognition of the writer remains a widely studied problem. To evaluate the effectiveness of our proposed approach, it is important to compare its performance with other well-established identification systems in the literature. A comparison of our system with those evaluated on the same datasets (CVL, BFL, IAM, QUWI-AR, and QUWI-EN) is presented in tables 4,5,6,7, and8.

TABLE 4. *Performance Comparison with State-of-the-art on the BFL database*

| System | Method | Number of writers | Top-1 |
|---|---|---|---|
| [8] | texture descriptors (LBP & LPQ) | 315 | 99.2 |
| [47] | Edge-hinge and Run-length | 315 | 98.4 |
| [20] | implicit shape codebook | 315 | 98.3 |
| [51] | LBP & oBIF | 315 | 98.6 |
| [52] | SIFT & SVM | 315 | 98.7 |
| **Our** | **Fragment Encoding (FAST/SIFT)** | **315** | **100** |

TABLE 5. *Performance Comparison with State-of-the-art on the CVL database*

| System | Method | Number of writers | Top-1 |
|---|---|---|---|
| [22] | Convolutional Neural Networks | 310 | 98.9 |
| [44] | BDCT descriptors | 310 | 99.6 |
| [23] | Convolutional Neural Networks | 310 | 99.2 |
| [47] | Edge-hinge and Run-length | 310 | 94.8 |
| [20] | shape codebook | 310 | 94.3 |
| [46] | LSTP | 310 | **100** |
| [27] | Convolutional Neural Networks | 310 | 99.1 |
| [53] | Convolutional Neural Networks | 310 | 99.7 |
| [54] | Residual Transformer | 310 | 93.3 |
| [55] | Diagonal gradient | 310 | 100 |
| [56] | Codebook | 300 | 99 |
| [57] | CNN | 310 | 82.1 |
| **Our** | **Fragment Encoding (SIFT)** | **310** | **100** |

Tables 4 and 5 demonstrate that our system based on FAST keypoint fragments for the BFL dataset and (SIFT/ SURF/ FAST/ Harris) fragments for the CVL dataset achieves the best performance with a $Top - 1$ identification rate of 100%. While in tables 6,7, and 8, although our system does not achieve the best performance on the QUWI-AR and QUWI-EN datasets, it is ranked second after our previous work, which is based on

TABLE 6. *Performance Comparison with State-of-the-art on the QUWI(AR) database*

| System | Method | Number of writers | Top-1 |
|---|---|---|---|
| [58] | direction, curvature, and tortuosity | 1017 | 70.1 |
| [16] | HOG and GLRL | 1017 | 76.3 |
| [28] | Convolutional Neural Networks | 1017 | 92.2 |
| [26] | Convolutional Neural Networks | 1017 | **99.8** |
| Our | Fragment Encoding (Harris) | 1017 | 96.9 |

TABLE 7. *Performance Comparison with State-of-the-art on the QUWI(EN) database*

| System | Method | Number of writers | Top-1 |
|---|---|---|---|
| [58] | direction, curvature, and tortuosity | 1017 | 70.1 |
| [16] | HOG and GLRL | 1017 | 76.3 |
| [28] | Convolutional Neural Networks | 1017 | 92.2 |
| [26] | Convolutional Neural Networks | 1017 | **99.7** |
| Our | Fragment Encoding (SURF) | 1017 | 98.3 |

TABLE 8. *Performance comparison with state-of-the-art on the IAM dataset*

| System | Method | Number of writers | Top-1 |
|---|---|---|---|
| [6] | Codebooks | 650 | 91.0 |
| [59] | SIFT | 657 | 98.5 |
| [60] | One-Class Classifier | 657 | 94.5 |
| [61] | SIFT and RootSIFT | 657 | 97.8 |
| [62] | Convolutional Neural Network | 657 | 93.1 |
| [27] | Convolutional Neural Network | 657 | 96.3 |
| [26] | Convolutional Neural Networks | 657 | **99.5** |
| [53] | Convolutional Neural Networks | 657 | 98.3 |
| [54] | Residual Transformer | 657 | 91.4 |
| [57] | CNN | 657 | 87.7 |
| Our | Fragment Encoding (FAST) | 657 | 97 |

a combination of deep learning and fragments extracted around the two keypoints FAST and Harris Corner Detector. However, our proposed technique remains simpler and faster than deep learning-based approaches, which require significant time for the training step. For instance, to train a ResNet-34 model on the QUWI-AR dataset, it took about a week in [26], while the fragment encoding-based method required only three hours.

5. **Conclusion.** This paper presents an innovative technique for the automatic identification of writers from handwritten documents. The proposed approach utilizes the existence of interesting pixels (keypoints) in handwritten images and the redundancy of certain patterns and handwriting gestures that are repeated for each writer through their handwriting. In contrast to classical methods that directly submit fragment information to a classifier, the proposed approach employs the VLAD encoding method to extract information from local features of these fragments, generating a global descriptor for each handwritten document to improve system performance. The technique is evaluated on five datasets of various languages and sizes, and the obtained results are comparable to the best classification rates achieved thus far.

The investigation of keypoint fragment encoding has the potential to be extended to various fields of machine vision, including signature verification and gender identification. Additionally, it may be worthwhile to explore the application of this technique in a multi-script environment, where the training and testing elements consist of handwritten documents written in multiple languages.

**Acknowledgment.** The authors gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

## REFERENCES

[1] R. P. Tett and C. A. Palmer, “The validity of handwriting elements in relation to self-report personality trait measures,” *Personality and individual differences*, vol. 22, no. 1, pp. 11–18, 1997.

[2] J. R. Beech and I. C. Mackintosh, “Do differences in sex hormones affect handwriting style? evidence from digit ratio and sex role identity as determinants of the sex of handwriting,” *Personality and individual differences*, vol. 39, no. 2, pp. 459–468, 2005.

[3] V. Christlein, M. Gropp, S. Fiel, and A. Maier, “Unsupervised feature learning for writer identification and writer retrieval,” in *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, vol. 1. IEEE, 2017, pp. 991–997.

[4] S. N. Srihari, S.-H. Cha, H. Arora, and S. Lee, “Individuality of handwriting,” Journal of *Forensic Sciences*, vol. 47, no. 4, pp. 856–872, 2002.

[5] M. Bulacu and L. Schomaker, “Text-independent writer identification and verification using textural and allographic features,” Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 29, no. 4, pp. 701–717, 2007.

[6] I. Siddiqi and N. Vincent, “Text independent writer recognition using redundant writing patterns with contour-based orientation and curvature features,” Pattern Recognition, vol. 43, no. 11, pp. 3853–3865, 2010.

[7] A. Semma, S. Lazrak, and Y. Hannad, “Enhancing writer identification with local gradient histogram analysis,” in The Proceedings of the International Conference on Smart City Applications. Springer, 2023, pp. 111–122.

[8] D. Bertolini, L. S. Oliveira, E. Justino, and R. Sabourin, “Texture-based descriptors for writer identification and verification,” Expert Systems with Applications, vol. 40, no. 6, pp. 2069–2080, 2013.

[9] D. Chawki and S.-M. Labiba, “A texture based approach for arabic writer identification and verification,” in 2010 International Conference on Machine and Web Intelligence. IEEE, 2010, pp. 115–120.

[10] V. Ojansivu and J. Heikkil¨a, “Blur insensitive texture classification using local phase quantization,” in International conference on image and signal processing. Springer, 2008, pp. 236–243.

[11] T. Ojala, M. Pietikainen, and T. Maenpaa, “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,” IEEE Transactions on pattern analysis and machine intelligence, vol. 24, no. 7, pp. 971–987, 2002.

[12] N. Bendaoud, Y. Hannad, A. Samaa, and M. E. Y. El Kettani, “Effect of the sub-graphemes’ size on the performance of off-line arabic writer identification,” in International Conference on Big Data, Cloud and Applications. Springer, 2018, pp. 512–522.

[13] M. Awais, M. J. Iqbal, I. Ahmad, M. O. Alassafi, R. Alghamdi, M. Basheri, and M. Waqas, “Real-time surveillance through face recognition using hog and feedforward neural networks,” IEEE Access, vol. 7, pp. 121 236–121 244, 2019.

[14] A. Semma, S. Lazrak, Y. Hannad, and M. E. Y. El Kettani, “Writer identification using vlad encoding of the histogram of gradient angle distribution,” in E3S Web of Conferences, vol. 351. EDP Sciences, 2022, p. 01073.

[15] Y. Hannad, I. Siddiqi, Y. El Merabet, and M. El Youssfi El Kettani, “Arabic writer identification system using the histogram of oriented gradients (hog) of handwritten fragments,” in Proceedings of the Mediterranean Conference on Pattern Recognition and Artificial Intelligence, 2016, pp. 98–102.

[16] Y. Hannad, I. Siddiqi, C. Djeddi, and M. E.-Y. El-Kettani, “Improving arabic writer identification using score-level fusion of textural descriptors,” IET Biometrics, vol. 8, no. 3, pp. 221–229, 2019.

[17] C. Djeddi, L.-S. Meslati, I. Siddiqi, A. Ennaji, H. El Abed, and A. Gattal, “Evaluation of texture features for offline arabic writer identification,” in 2014 11th IAPR international workshop on document analysis systems. IEEE, 2014, pp. 106–110.

[18] I. Siddiqi and N. Vincent, “Writer identification in handwritten documents,” in Ninth International Conference on Document Analysis and Recognition (ICDAR 2007), vol. 1. IEEE, 2007, pp. 108–112.

[19] C. Djeddi and L. Souci-Meslati, “Une approche locale en mode ind´ependant du texte pour l’identification de scripteurs: Application ‘a l’´ecriture arabe,” in Colloque international francophone sur l’ecrit et le document. Groupe de Recherche en Communication Ecrite, 2008, pp. 151–156.

[20] A. Bennour, C. Djeddi, A. Gattal, I. Siddiqi, and T. Mekhaznia, “Handwriting based writer recognition using implicit shape codebook,” Forensic science international, vol. 301, pp. 91–100, 2019.

[21] A. Semma, Y. Hannad, and M. E. Y. El Kettani, “Impact of the cnn patch size in the writer identification,” in Networking, Intelligent Systems and Security. Springer, 2022, pp. 103–114.

[22] S. Fiel and R. Sablatnig, “Writer identification and retrieval using a convolutional neural network,” in International Conference on Computer Analysis of Images and Patterns. Springer, 2015, pp. 26–37.

[23] V. Christlein and A. Maier, “Encoding cnn activations for writer recognition,” in 2018 13th IAPR International Workshop on Document Analysis Systems (DAS). IEEE, 2018, pp. 169–174.

[24] S. He and L. Schomaker, “Gr-rnn: Global-context residual recurrent neural networks for writer identification,” Pattern Recognition, p. 107975, 2021.

[25] A. Semma, S. Lazrak, Y. Hannad, M. Boukhani, and Y. El Kettani, “Writer identification: The effect of image resizing on cnn performance,” The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 46, pp. 501–507, 2021.

[26] A. Semma, Y. Hannad, I. Siddiqi, C. Djeddi, and M. E. Y. El Kettani, “Writer identification using deep learning with fast keypoints and harris corner detector,” Expert Systems with Applications, vol. 184, p. 115473, 2021.

[27] S. He and L. Schomaker, “Fragnet: Writer identification using deep fragment networks,” IEEE Transactions on Information Forensics and Security, vol. 15, pp. 3013–3022, 2020.

[28] A. Rehman, S. Naz, M. I. Razzak, and I. A. Hameed, “Automatic visual features for writer identification: A deep learning approach,” IEEE access, vol. 7, pp. 17 149–17 157, 2019.

[29] A. Semma, Y. Hannad, I. Siddiqi, S. Lazrak, and M. E. Y. E. Kettani, “Feature learning and encoding for multi-script writer identification,” International Journal on Document Analysis and Recognition (IJDAR), feb 2022.

[30] Y. Hannad, I. Siddiqi, and M. E. Y. El Kettani, “Writer identification using texture descriptors of handwritten fragments,” Expert Systems with Applications, vol. 47, pp. 14–22, 2016.

[31] D. G. Lowe, “Object recognition from local scale-invariant features,” in Proceedings of the seventh IEEE international conference on computer vision, vol. 2. Ieee, 1999, pp. 1150–1157.

[32] H. Bay, T. Tuytelaars, and L. Van Gool, “Surf: Speeded up robust features,” in European conference on computer vision. Springer, 2006, pp. 404–417.

[33] C. G. Harris, M. Stephens et al., “A combined corner and edge detector.” in Alvey vision conference, vol. 15, no. 50. Citeseer, 1988, pp. 10–5244.

[34] E. Rosten and T. Drummond, “Machine learning for high-speed corner detection,” in European conference on computer vision. Springer, 2006, pp. 430–443.

[35] H. Jegou, F. Perronnin, M. Douze, J. S´anchez, P. Perez, and C. Schmid, “Aggregating local image descriptors into compact codes,” IEEE transactions on pattern analysis and machine intelligence, vol. 34, no. 9, pp. 1704–1716, 2011.

[36] F. Perronnin and C. Dance, “Fisher kernels on visual vocabularies for image categorization,” in 2007 IEEE conference on computer vision and pattern recognition. IEEE, 2007, pp. 1–8.

[37] Y. Wang, Y.-P. Huang, and X.-J. Shen, “St-vlad: Video face recognition based on aggregated local spatial-temporal descriptors,” IEEE Access, vol. 9, pp. 31 170–31 178, 2021.

[38] A. Vinay, V. S. Shekhar, C. A. Kumar, A. S. Rao, G. R. Shenoy, K. B. Murthy, and S. Natarajan, “Face recognition using vlad and its variants,” in Proceedings of the Sixth International Conference on Computer and Communication Technology 2015, 2015, pp. 233–238.

[39] S. M. Omohundro, Five balltree construction algorithms. International Computer Science Institute Berkeley, 1989.

[40] U.-V. Marti and H. Bunke, “The iam-database: an english sentence database for offline handwriting recognition,” International Journal on Document Analysis and Recognition, vol. 5, no. 1, pp. 39–46, 2002.

[41] S. Al Maadeed, W. Ayouby, A. Hassa¨ıne, and J. M. Aljaam, “Quwi: An arabic and english handwriting dataset for offline writer identification,” in 2012 International Conference on Frontiers in Handwriting Recognition. IEEE, 2012, pp. 746–751.

[42] C. Freitas, L. S. Oliveira, R. Sabourin, and F. Bortolozzi, “Brazilian forensic letter database,” in 11th International workshop on frontiers on handwriting recognition, Montreal, Canada, 2008.

[43] F. Kleber, S. Fiel, M. Diem, and R. Sablatnig, "Cvl-database: An off-line database for writer retrieval, writer identification and word spotting," in 2013 12th international conference on document analysis and recognition. IEEE, 2013, pp. 560–564.
[44] F. A. Khan, M. A. Tahir, F. Khelifi, A. Bouridane, and R. Almotaeryi, "Robust off-line text independent writer identification using bagged discrete cosine transform features," Expert Systems with Applications, vol. 71, pp. 404–415, 2017.
[45] A. Chahi, Y. Ruichek, R. Touahni et al., "An effective and conceptually simple feature representation for off-line text-independent writer identification," Expert Systems with Applications, vol. 123, pp. 357–376, 2019.
[46] A. Chahi, Y. Ruichek, R. Touahni et al., "Local gradient full-scale transform patterns based off-line text-independent writer identification," Applied Soft Computing, p. 106277, 2020.
[47] Y. Kessentini, S. BenAbderrahim, and C. Djeddi, "Evidential combination of svm classifiers for writer recognition," Neurocomputing, vol. 313, pp. 1–13, 2018.
[48] J. Delhumeau, P.-H. Gosselin, H. J ´egou, and P. P ´erez, "Revisiting the vlad image representation," in Proceedings of the 21st ACM international conference on Multimedia, 2013, pp. 653–656.
[49] R. Arandjelovic and A. Zisserman, "All about vlad," in Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, 2013, pp. 1578–1585.
[50] E. Spyromitros-Xioufis, S. Papadopoulos, I. Kompatsiaris, G. Tsoumakas, and I. Vlahavas, "An empirical study on the combination of surf features with vlad vectors for image search," in 2012 13th International Workshop on Image Analysis for Multimedia Interactive Services. IEEE, 2012, pp. 1–4.
[51] F. Abbas, A. Gattal, C. Djeddi, I. Siddiqi, A. Bensefia, and K. Saoudi, "Texture feature column scheme for single-and multi-script writer identification," IET Biometrics, vol. 10, no. 2, pp. 179–193, 2021.
[52] J. P. L. Sanchez, "Identificac¸ ˜ao de escritor por transformada sift e svm-linear na l´ıngua portuguesa," 2024.
[53] S. N. M. Khosroshahi, S. N. Razavi, A. B. Sangar, and K. Majidzadeh, "Deep neural networks-based offline writer identification using heterogeneous handwriting data: an evaluation via a novel standard dataset," Journal of Ambient Intelligence and Humanized Computing, pp. 1–20, 2022.
[54] P. Zhang, "Rstc: A new residual swin transformer for offline word-level writer identification," IEEE Access, vol. 10, pp. 57 452–57 460, 2022.
[55] A. Semma, S. Lazrak, M. Boukhani, and Y. Hannad, "Offline writer identification based on diagonal gradient angle of small fragments," in International Conference on Artificial Intelligence and Green Computing. Springer, 2023, pp. 92–105.
[56] M. N. Abdi and M. Khemakhem, "Offline text-independent arabic and chinese writer identification using a multi-segmentation codebook-based strategy." in ICPRAM, 2024, pp. 613–619.
[57] V. Kumar and S. Sundaram, "Utilization of information from cnn feature maps for offline word-level writer identification," Expert Systems with Applications, vol. 238, p. 121709, 2024.
[58] S. Al-Maadeed, A. Hassaine, A. Bouridane, and M. A. Tahir, "Novel geometric features for off-line writer identification," Pattern Analysis and Applications, vol. 19, no. 3, pp. 699–708, 2016.
[59] X. Wu, Y. Tang, and W. Bu, "Offline text-independent writer identification based on scale invariant feature transform," IEEE Transactions on Information Forensics and Security, vol. 9, no. 3, pp. 526–536, 2014.
[60] B. Hadjadji and Y. Chibani, "Two combination stages of clustered one-class classifiers for writer identification from text fragments," Pattern Recognition, vol. 82, pp. 147–162, 2018.
[61] F. A. Khan, F. Khelifi, M. A. Tahir, and A. Bouridane, "Dissimilarity gaussian mixture models for efficient offline handwritten text-independent identification using sift and rootsift descriptors," IEEE Transactions on Information Forensics and Security, vol. 14, no. 2, pp. 289–303, 2018.
[62] H. T. Nguyen, C. T. Nguyen, T. Ino, B. Indurkhya, and M. Nakagawa, "Text-independent writer identification using convolutional neural network," Pattern Recognition Letters, vol. 121, pp. 104–112, 2019.