Bulletin of Electrical Engineering and Informatics
Vol. 13, No. 6, December 2024, pp. 4157~4167
ISSN: 2302-9285, DOI: 10.11591/eei.v13i6.7906

Handwritten digit recognition using a column scheme-based local directional number pattern

Mohammed Aouine[1,2], Abdeljalil Gattal[2], Chawki Djeddi[2], Faycel Abbas[2]
[1]Laboratoire D'informatique et Mathématiques, Department of Computer Science, Akli Mohand Oulhadj University, Bouira, Algeria
[2]Laboratoire de Vision et d'Intelligence Artificielle (LAVIA), Université Larbi Tébessi, Tébessa, Algeria

Article Info

Article history:

Received Nov 20, 2023
Revised Apr 6, 2024
Accepted May 17, 2024

Keywords:

Column scheme
CVL dataset
Handwritten digit recognition
Local directional number pattern
Support vector machines
Texture feature

ABSTRACT

One of the most well-known challenges in computer vision and machine learning is the recognition of handwritten digits. This study presents an advanced approach to improving isolated-digit recognition through the use of advanced feature extraction techniques. For example, digit recognition is commonly used to read numbers on forms and checks in banks. This paper introduces a novel method of extending the local directional number pattern (LDNP) to a column scheme using two different masks and their resolutions. A new descriptor of the LDNP column scheme is being proposed that combines derivative Gaussian and Kirsch masks in order to enhance textural analysis and capture more detailed local textual information. This approach is highly efficient and robust, able to handle variations in size, shape, and slant. Additionally, the support vector machine (SVM) is employed as a classifier, which has been shown to make better decisions. The empirical investigation is carried out using the CVL dataset, resulting in recognition rates that are comparable with the latest advancements in the field. The overall precision of 96.64% is achieved, outperforming existing similar works.

Corresponding Author:

Mohammed Aouine
Laboratoire D'informatique et Mathématiques, Department of Computer Science
Akli Mohand Oulhadj University
Bouira, Algeria
Email: Mohamed.aouine@univ-tebessa.dz

1. INTRODUCTION

Automated recognition of handwritten digits poses significant challenges in traditional pattern recognition, with diverse applications such as postal code identification, bank check processing, and digit-based data storage. The primary difficulties in understanding handwritten digits stem from variations in shape, size, inclination, and particularly, individual writing styles, rendering the task notably intricate. Common approaches typically include feature extraction and classification techniques. In the realm of handwritten digit recognition, traditional machine learning methods primarily focus on these two aspects [1]–[6]. Several classification methods have been explored, including hidden Markov models (HMM) [2]–[4], k-nearest neighbors (KNN) [3], [6], neural networks [1], [2], and support vector machines (SVM) [1], [6]–[9]. These methods have shown higher performance compared to their modern counterparts, but they may struggle with robustness and computational complexity.

Among the first kinds, feature extraction methods usually use three categories of features: statistical, structural, and textural. Statistical features are gathered directly from the statistical information distribution of input images [5], [7], [8]. Structural features use the topological and geometric properties [5], [7], [8], and textural features [8]–[10] are used for identifying objects or zones of interest in an image. Furthermore, some methods are also available that are based on the combination of numerous feature types [7], [11], [12]. In the literature, there are several approaches that use local features that appear just locally and perform better in terms of boosting robustness against outside interference. These descriptors are a type of detailed information that can be displayed consistently and clearly, such as texture information. Among these methods, oriented basic image features (oBIFs) [13], local phase quantization (LPQ) [12], [14], local binary pattern (LBP) [15]–[17], and oBIFs column histogram [18], [19] are the most popular and have demonstrated enhanced robustness against external interference. However, they may still face limitations in capturing all relevant features effectively. Many advanced methods have been developed to overcome the shortcomings of LBP. For example, the local directional number pattern (LDNP) [20], which encodes the directional information in the neighborhood, has previously been used for face analysis. The method can solve the problems of random noise and non-monotonic illumination variation [20].

While previous studies explored the capabilities of methods like KNN, HMM, and SVM, they often overlooked the potential of local feature descriptors such as oBIFs, LBP, and LPQ in enhancing robustness against external interference. However, despite their advantages, these methods may fail to capture all relevant features effectively, particularly at different resolutions. Deep learning has made advances in the last few years and achieved promising results in lots of applications, including handwritten recognition [2], [21]. Unlike traditional methods, these methods do not need a specific step to extract the crucial features. One of the currently well-known architectures proposed for digit recognition using deep learning is deep convolutional neural network (CNN). The deep CNN algorithms consist of a feature extractor network, and its fully connected layers are used for estimating the probabilities of handwritten digits.

Despite their success, deep learning methods have limitations that include the need for a significant amount of computing power and a large database of normalized images. In a study by Bendib *et al.* [21], utilized a deep CNN with various configurations from the CNN architecture to identify digits from images in the CVL database. The CNN architecture comprised three convolutional layers, followed by three max pooling layers, and concluded with a fully connected layer. Through study in this area, the goal is to develop advanced feature extraction methods that improve recognition rates, particularly on datasets like CVL used in competitions. This research aims to contribute to the advancement of automated handwritten digit recognition systems, enhancing their accuracy, robustness, and applicability in real-world scenarios. To address these limitations, the LDNP column scheme method has been proposed. LDNP encodes both the intensity variations and structural information of local textural patterns. The initial idea of LDNP is to use a single mask to capture both the structural information and intensity variations of local textural patterns. However, a major limitation of this method is that it may not effectively capture all of the main features at different resolutions. This issue can be mitigated by employing the LDNP column scheme, which utilizes various masks and resolutions to capture features that may not be adequately represented by a single mask, amalgamating them to enhance the encoded information. This approach uses information about the entire neighborhood to distinguish the intensity of the texture. Developing feature extraction can increase the recognition rate of the CVL dataset [6]. This point has been investigated. The handwritten digit recognition competition (HDRC), which was organized in collaboration with ICDAR 2013, makes use of this dataset [6].

The structure of the paper is as follows: the related work is briefly discussed in section 2. In section 3, we will present the computation of the LDNP column scheme that we employed to characterize the handwritten digit and a preview of a SVM classifier. In section 4, we discuss the experimental settings, the realized results, and system stability. Finally, we present the concluding remarks about the consequences of the results obtained in section 5.

2. RELATED WORKS

Offline handwritten digit recognition systems focused on feature extraction methods have been researched over the last few decades for many applications, including the recognition of postal codes, bank check amounts, and storing handwriting. The offline handwritten digit recognition issue focuses on enhancing performance by optimizing the feature extraction that can separate between digits, the most significant module of any pattern recognition system. Gattal *et al.* [7] conducted a notable investigation on unnormalized digits within the CVL dataset, exploring various combinations of global, local, structural, and statistical features for handwritten digit recognition. The features assessed encompass Zernike moments, Hu’s moment invariants, profile and project-based features, skew angle, background and foreground features, as well as Ridgelet transforms derived from different regions across the digits. In a separate work, Gattal *et al.* [9] demonstrated that the amalgamation of oriented oBIF with background concavity features significantly enhances the accuracy of handwritten digit recognition, avoiding the need for size normalization.

However, Gattal and Abbas [12], introduced a texture-based encoding approach utilizing LPQ and LBP to attain significant recognition rates for unnormalized isolated handwritten digits. In a recent investigation by Gattal *et al.* [19], they introduced a novel method to broaden the oBIFs column to encompass multiple scale features employing SVM classifiers. The authors demonstrated how combining oBIF images for scale parameters at two or more scales within a column scheme can yield optimal performance. Additionally, their system attained satisfactory recognition results for unnormalized digits on the CVL dataset, achieving nearly 96% accuracy. A summary outlining prominent contributions to both normalized and unnormalized CVL datasets from existing literature is given in the accompanying Table 1.

Table 1. Performance comparison of well-known offline handwritten digit recognition systems on CVL datasets

Rank	Study	Classifier	Precision (%)	Normalized digits
1	Salzburg [6]	FIRMLP	97.74	Yes
2	Deep CNN [21]	CNN	96.63	Yes
3	Gattal *et al*. 2014 [7]	SVM	96.62	No
4	Multi-scale oBIF column [19]	SVM	95.84	No
5	LPQ and LBP features [12]	SVM	95.36	No
6	oBIFs and background features [9]	SVM	95.21	No
7	Jadavpur [6]	SVM	94.75	No
8	Tébessa I [6]	SVM	77.53	No

Recent advancements have led to a significant shift towards the utilization of deep learning and machine learning techniques to enhance recognition capabilities. In the paper of [22], Agrawal propose a new model for handwritten digit and character recognition using deep learning techniques, incorporating ResNet 151 for feature extraction and an enhanced Bi-LSTM-DNN architecture. In another recent study, the work of [23] analyzed the impact of the hybrid convolutional vision transformer (ViT) model on the capacity to recognize handwritten digits. However, realtime data often includes distortions, noise, and a variety of writing styles. To address this, the study of [24] presents the results of their study on handwritten digit recognition applying state-of-the-art feature extraction and classification approaches using well-known image databases. The authors incorporated ten feature vectors and eight classifiers, resulting in improved performance. In another study, Abdelmoumene *et al*. [25] present a novel approach to data augmentation, employing varied encrypted digit images and deep learning-based image fusion techniques to improve the efficacy of handwritten digit recognition. This literature review highlights the evolution from traditional feature extraction methods to the latest deep learning models, which have set new benchmarks in the accuracy and efficiency of handwritten digit recognition systems. This transition marks a move towards more adaptive and intelligent systems capable of handling the complexities of handwritten digit recognition with unprecedented precision.

3. PROPOSED APPROACH

LDNP method has been successfully used in various areas such as texture classification [20], face expression analysis [26], and face recognition [27]. Our goal in using this feature is to complement the LDNP column scheme and improve the representation of handwritten digits. These methods offer advanced techniques for extracting features that can enhance the accuracy of handwritten digit recognition systems. However, the LDNP method has limitations in its ability to capture essential features using a single mask and resolution. To address this limitation, we propose the LDNP column scheme, which utilizes multiple masks and resolutions to extract features that may be overlooked by a singular approach. This problem can be alleviated by using the LDNP column scheme that we have discussed in the following section. As a result, our goal is to develop a system that can reliably recognize handwritten digits. Figure 1 shows a schematic illustration of our proposed approach.

3.1. Local directional number pattern

Feature extraction plays a crucial role in image classification, allowing the mapping of images into points within a feature space where images of the same class tend to cluster together. In our research, we decided to capture the contour, curvature, and texture details of handwriting to describe the hand. This involves computational features such as the LDNP and it is column scheme, which are discussed next.

The LDNP represents the texture pattern as well as its intensity transitions, encoded as a binary code of six bits corresponding to each pixel of an image. Therefore, the idea of LDNP is to use the compass mask to identify the edge response of the neighborhood and get the top directional numbers and up values to create the pattern. The positive and negative directions of the edge responses indicate the gradient direction of dark and bright areas in the neighborhood and thus provide useful information about neighboring structures. Consequently, this discrimination between dark and bright responses allows the LDNP to distinguish between blocks with swapped or reversed positive and negative directions. Moreover, these transitions often occur on the handwritten digit. For example, the nine numbers have transitions of different strengths at the top and bottom. It is therefore important to distinguish between them. Furthermore, an LDNP has been proposed to avoid the drawbacks of previous methods, such as the local directional pattern (LDiP) [26] and local derivative pattern (LDeP) [27]. The last method treats all directions equally but loses some orientation information. Likewise, they are susceptible to illumination variations and noise. To circumvent these problems, the LDNP uses the sign and form of directional numbers to improve the encoded structural information according to two different masks: a derivative-Gaussian mask and a Kirsch compass mask (KCM). After applying the filter method, LDNP codes the top two numbers from each neighborhood into a single number for the selection of important information about each pixel's neighborhood.

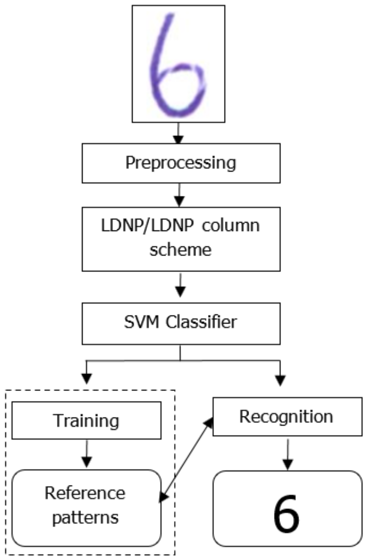

Figure 1. An overview of the proposed approach

3.2. LDNP column scheme

The main limitation of the LDNP method is that it fails to capture all of the main features using only one mask and resolution. This issue can be addressed by implementing the LDNP column scheme, which utilizes multiple masks and resolutions to gather features that may be mistreated by a single mask. By combining these multiple masks, the encoded information can be extended and improved. To produce the LDNP column scheme, we must apply a compass mask to generate edge responses through two various asymmetric masks, derivative-Gaussian and Kirsch as illustrated in Figures 2 and 3. These masks work in gradient space to reveal handwritten digit structure. In addition, we use a derivative-Gaussian mask by applying Gaussian smoothing in order to make the method robust in the presence of noise.

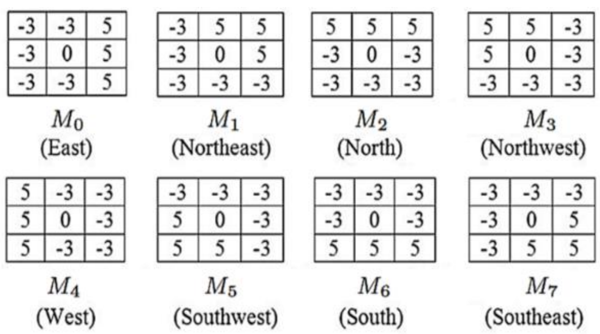

Figure 2. Representation of a 3×3 KCM

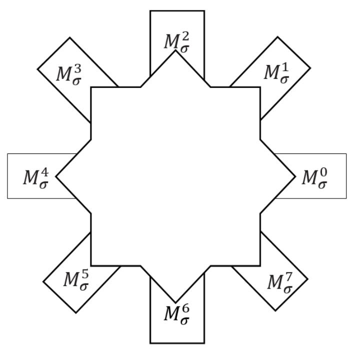

Figure 3. Derivative of Gaussian compass masks

3.2.1. Kirsch compass mask

KCM [28] is also a derivative mask used for edge detection. This method finds edges in all eight compass directions. KCMs modify the mask to suit the requirements of the input image. In the LDNP method, the KCM is rotated 450 in order to get the edge response in eight distinct directions. We can find edges in the following eight directions, as shown in Figure 2.

The LDNP method represents the edge response by analyzing each mask, $M_0, \dots, M_7$, and combining the main directional numbers to extract the LDNP features. Edge responses are not identically significant; hence, the presence of high positive or negative values reveals significant dark or bright areas. Therefore, the three higher significant bits (representing the top positive directional number) together with the three lower significant bits (representing the top negative directional number) of the code are identified as significant regions, as illustrated in Figure 2. The LDNP can be expressed in (1):

$$LDNP(x,y) = 8 \times \arg_i \max\{\ conv_i(x,y)\big|\ 0 \le i \le 7\ \} + \arg_j \min\{\ conv_j(x,y)\big|0 \le\ j \le 7\ \})(1)$$

where *(x, y)* represents the central pixel of the encoded neighborhood, $conv_i$ is the convolution of the original image I, and the M_i is a mask, defined by (2):

$$conv_i = I\ *\ M_i \qquad (2)$$

These convolution masks typically operate in small neighborhoods ranging from (3×3) to (11×11). In our study, the (11×11) KCM which can detect more edges in different directions has been considered for enhanced edge detection.

3.2.2. Derivative-Gaussian compass mask

To generate the LDNP code in the gradient space, we use a derivative-Gaussian compass mask to calculate the edge responses. This mask is inspired by the Kirsch mask [28], but instead uses the derivative of a skewed Gaussian to create an asymmetric compass mask. This allows for the computation of edge responses for smoothed handwritten digits, as shown in Figures 2 and 3. The mask is designed to be robust against noise and changes in illumination, while still producing strong edge responses [20]. The smoothed function is obtained by convolving the original image I with the Gaussian weight function G^{σ}. The Gaussian mask is defined as (3):

$$G^{\sigma}(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \qquad (3)$$

where (x, y) are the spatial locations and σ is the Gaussian bell width.
The mask M_{σ} can be defined as (4):

$$M_{\sigma}\ (x\ ,y) = G'^{\sigma}(x+k,y) *\ G^{\sigma}\ (x\ ,y) \qquad (4)$$

where G'^{σ}represents the derivative of G^{σ} with respect to x, $*$ denotes the convolution operation, and k is the Gaussian offset from its center. For this offset, we have used half of the mask diameter in our work.

In Figure 3, eight various directions generate a compass mask $M_{\sigma}^0, M_{\sigma}^1, M_{\sigma}^2, M_{\sigma}^3, M_{\sigma}^4, M_{\sigma}^5, M_{\sigma}^6, M_{\sigma}^7$ by rotating them 45° apart. Thus, define the code produced through this mask like this:

$$LDNP_{\sigma}^{G}(x,y) = 8 \times arg_i ma\,x\, conv_i(x,y)|0 \le i \le 7 + arg_j mi\,n\, conv_j(x,y)|0 \le j \le 7 \quad (5)$$

where $conv_i$ denotes the convolution of the original image *I*, defined by give the equation numbers in the texts for all throughout the entire manuscript:

$$conv_i = I\ *\ M_{\sigma}^{i} \qquad (6)$$

The LDNP column scheme enriching the LDNP feature representation that produces by crossing two different masks and its resolutions (derivative-Gaussian and a KCM) to take advantage of the local textural variations and the intensity variations between the two masks as shown in Figure 4.

A vector of 56 directional features defines the dimension of the LDNP features set. As such, we have identified and included 56 items in the LDNP dictionary. To enhance the performance of our isolated handwritten digit recognition system, we decided to combine the LDNP image with a derivative-Gaussian compass mask and a KCM at each position to create LDNP column features. This results in an increase in the

LDNP column scheme histogram to 3,136, with a size of 56×56. Figure 4 illustrates the different stages of the LDNP column design. In Figure 4(a), a binarized image is shown using Sauvola's method [29]. Figure 4(b) displays an LDNP image with a derivative-gaussian compass mask, while Figure 4(c) shows an LDNP image with a KCM. Figure 4(d) displays the crossing of two LDNP images by examining pairs of two preceding LDNP images at each point, resulting in an LDNP column at each position. Finally, Figure 4(e) presents the histogram.

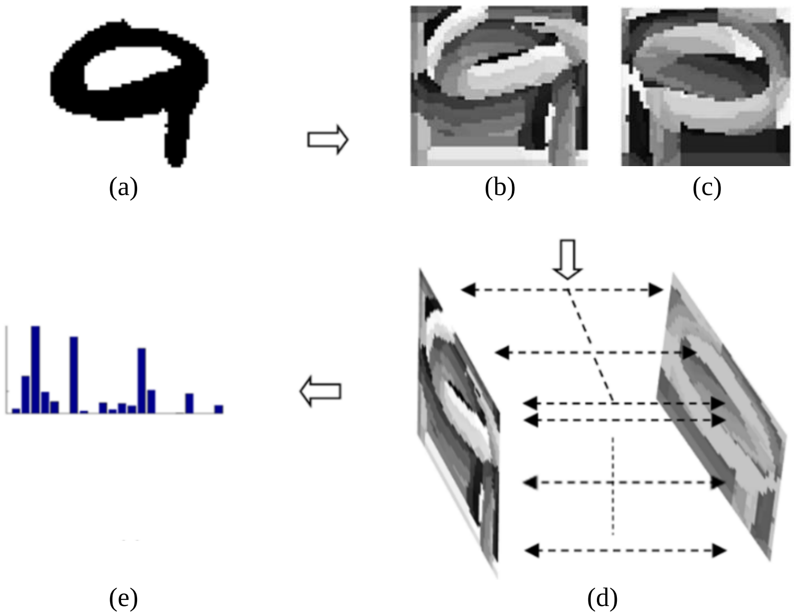

Figure 4. Various steps involved in the LDNP column scheme; (a) the binarized image, (b) the LDNP image using a derivative-gaussian compass mask, (c) the LDNP image using a KCM, (d) the two LDNP images are crossed to create an LDNP column at each location, and (e) the histogram is computed

In summary, our proposed method, the LDNP column scheme, is founded on directional numbers instead of bit strings. These numbers encode information about the neighborhood in the derivative-Gaussian compass mask and the KCM. This approach allows for more information to be encoded and better discrimination of textures, while also being robust against illumination changes and noise. However, since the LDNP column scheme only encodes certain patterns without location information in the whole image (WI), we use the uniform grid sampling (UGS) method to aggregate location information for the descriptor [30]. This involves dividing the image into smaller regions and extracting a histogram from each region. The last method performed on the digit image involves producing rectangular regions for sampling, with each region being of similar size and shape. The extracted feature vector is then standardized to have a zero mean and unit variance. Finally, classification is performed using a SVM with a one-against-all implementation [31]. This involves building 10 binary classifiers to address a 10-class digit recognition problem, with each classifier trained to discriminate one class from all others. The classifiers are trained using the proposed features extracted from the handwritten-digit image, with the class whose decision function yields the maximum value being selected.

$$Max\{f_j(x); j = 0,..9\} \tag{7}$$

where x denotes the feature vector of the query handwritten digit and f refers to decision function. Two crucial parameters essential for training an SVM are the regularization parameter (C) and the radial basis function (RBF) with kernel parameter (σ). The RBF is a popular kernel function used with SVM. It is defined as (8):

$$K(x, x') = exp\left(-\frac{\|x-x'\|^2}{2\sigma^2}\right) \tag{8}$$

where $\|x - x'\|^2$ is the Euclidean distance between the input data points x and x', and σ is a free parameter. The following section presents the experimental setup and, on the other hand, the corresponding results.

4. EXPERIMENTS AND RESULTS

The studies were conducted using the CVL single digit database [6], which has three sets of numbers: 7,000 for training, the same number for validation, and 21,780 for evaluation. Prior to extracting features, these digit images were binarized using Sauvola's method [29]. The features were then extracted from the binary digit images without using size normalization. A sequence of experiments were performed to evaluate the effectiveness of various methods, including texture-based encoding based on the proposed LDNP column scheme, LDNP code using derivative-Gaussian compass mask (LDNP-DGCM), LDNP-KCM, and others such as LBP, LPQ, oBIFs, and oBIFs column histogram [19]. These methods were generated using optimal parameter values to achieve the best performance from the WI. The performance of the system was measured using the standard recall measure, computed in a similar manner to the ICDAR 2013-digit recognition competition [6]. The results of the experiments are summarized in Table 2.

Table 2. Performance of different features

Features	Optimal parameters	Dimension	Recall (%)
LPQ	$w = 13 \times 13$	256	87.90
	$w = 15 \times 15$	256	87.72
	$w = 11 \times 11$	256	87.20
LBP	$P = 16, R = 4$	243	87.75
	$P = 8, R = 4$	59	87.10
	$P = 16, R = 2$	243	81.57
oBIFs	$\sigma = 4, \varepsilon = 0.1$	23	76.53
	$\sigma = 8, \varepsilon = 0.1$	23	72.29
	$\sigma = 2, \varepsilon = 0.1$	23	68.05
LBP column scheme	$[P = 16, R = 4]$ $[P = 16, R = 2]$	59,049	87.69
	$[P = 8, R = 4]$ $[P = 8, R = 2]$	3,481	87.10
	$[P = 16, R = 4]$ $[P = 8, R = 4]$	14,337	86.42
Proposed features			
LDNP-KCM	11×11	56	83.63
	13×13	56	83.14
	9×9	56	81.89
LDNP-DGCM	$\sigma = 2.3$	56	58.60
	$\sigma = 2.2$	56	85.34
	$\sigma = 2.5$	56	85.34
LDNP column scheme	$11 \times 11, \sigma = 2.3$	3,136	90.03
	$11 \times 11, \sigma = 2.2$	3,136	89.99
	$13 \times 13, \sigma = 2.3$	3,136	89.25

Table 2 presents the top three performance efficiencies of each texture-based encoding method, as well as our proposed features for digit recognition. It is evident that there is significant variation in the performance of these methods. For example, when using LBP, higher values of P=16 and radius R=4 generally result in improved recall rates. However, the use of LPQ with a window size of 13 leads to even higher recall rates, highlighting the importance of capturing spatial information across larger regions of the image. Similarly, the choice of parameters σ and ε significantly affects the performance of oBIFs. However, it is worth noting that performance does not consistently improve with different values of σ and only achieves a modest result of 76.53%.

In contrast, the LBP column scheme shows varying recall rates, with certain combinations outperforming others. On the other hand, our proposed LDNP column scheme, which is extracted from the WI using a derivative-Gaussian compass mask and KCM, shows promising results with recall rates reaching up to 90.03%. These findings suggest that a combination of feature extraction methods, such as the LDNP column scheme, yields superior performance in digit recognition tasks compared to individual methods. However, it is important to note that performance is highly sensitive to parameter selection, emphasizing the importance of tuning parameters for optimal results. These findings align with the objectives of the study, which aimed to explore different feature extraction techniques for digit recognition and compare their performance. However, it is worth mentioning that the study may have limitations, such as a lack of analysis of computational efficiency and robustness to noise and variations in input data.

In order to increase the effectiveness of the approach, we equally assessed the proposed method using both the WI and different image regions sampled through UGS. Table 3 summarizes the recall of the proposed method using the UGS method with and without the WI. Using the combination method scheme, the highest recall rate of 95.84% was obtained when a feature vector with a dimension of 560 was extracted from both the WI and a 3×3 grid.

Table 3. Recall results on the proposed method using UGS with and without the WI

Concatenate with WI	UGS size	LDNP-KCM		LDNP-DGCM		LDNP column scheme	
		Dimension	Recall (%)	Dimension	Recall (%)	Dimension	Recall (%)
No	1×2 grid	112	90.33	112	91.78	6,272	91.01
	2×1 grid	112	90.79	112	93.1	6,272	92.96
	2×2 grid	224	93.95	224	95.06	12,544	93.58
	1×3 grid	168	91.14	168	92.69	9,408	91.24
	3×1 grid	168	93.29	168	94.44	9,408	94.08
	2×3 grid	336	94.09	336	95.2	18,816	93.26
	3×2 grid	336	94.35	336	95.66	18,816	94.4
	3×3 grid	504	94.25	504	95.76	28,224	94.4
Yes	1×2 grid	168	90.97	168	91.88	9,408	92.78
	2×1 grid	168	90.97	168	93.18	9,408	93.13
	2×2 grid	280	94.37	280	95.08	15,680	94.27
	1×3 grid	224	91.79	224	92.77	12,544	92.43
	3×1 grid	224	93.55	224	94.45	12,544	94.13
	2×3 grid	392	94.24	392	95.23	21,952	93.87
	3×2 grid	392	94.65	392	95.73	21,952	94.78
	3×3 grid	560	94.51	560	95.84	31,360	94.48

Besides distinct features, we have also evaluated combinations of features to increase overall performance. In other words, we have transformed handwritten digits using a single-feature histogram into a multi-feature histogram. The four features that will be concatenated are the top-performing features that were taken from various image areas and the WI. These features include LDNP-KCM with a 3×2 grid, referred to as KCM32; LDNP-KCM with a 3×3 grid, referred to as KCM33; LDNP-DGCM with a 3×2 grid, referred to as DGCM32; LDNP-DGCM with a 3×3 grid, referred to as DGCM33; LDNP column scheme with a 3×2 grid, referred to as LDNPCS32; and LDNP column scheme extracted from the WI, referred to as LDNPCS11. Table 4 summarizes the performance of these combinations. The highest recall of 96.59% is reached when combining the DGCM33, KCM32, and LDNPCS11, creating a feature vector with a dimension of 4,088.

Table 4. Performance on different features combination

Features combinations	Dimension	Recall (%)
DGCM32 and KCM32	784	95.83
DGCM32 and KCM33	952	95.69
DGCM33 and KCM32	952	96.51
DGCM33 and KCM33	1,120	95.87
DGCM33 and LDNPCS32	22,512	95.78
DGCM33, KCM32, and LDNPCS32	22,904	95.15
DGCM32, KCM33, and LDNPCS32	22,904	95.48
DGCM33, KCM32, and LDNPCS11	4,088	96.59
DGCM33 and LDNPCS11	3,696	95.83

We have discovered that by combining local shape and contour information, the relationship between edge responses, and information from various textures, we are able to more accurately characterize digits. To further evaluate the effectiveness of our approach, we calculated the recall for each digit class individually in order to identify the most challenging classes for recognition. For these experiments, Table 5 displays the class-wise recognition rates. Overall, recall values are uniform across distinct digit classes. However, there are some digits (3, 8, and 9) that have a relatively low recall. It is worth noting that certain pairs, such as (‘1’, ‘9’), (‘8’, ‘9’), and (‘4’, ‘7’), present a more difficult recognition issue due to their low inter-class variation.

We evaluated our proposed approach against the leading digit recognition approaches that were submitted to the HDRC during ICDAR 2013. As a whole, seven teams submitted nine various systems to HDRC 2013, with only two systems not requiring size normalization (Jadavpur and Tébessa I). To ensure a fair comparison, we utilized the same database (CVL) and scoring protocol as the competition, as depicted in

Table 6. The results presented in Table 6 illustrate that our proposed approach achieves a precision level of 96.64%, which aligns closely with the performance of the top-ranking participant in the competition. It's important to highlight that our system directly extracts features from unnormalized images, allowing the capture of local textural patterns. These findings demonstrate the efficacy of the LDNP and its column scheme, coupled with a straightforward combination method, in accurately recognizing isolated digits.

Table 5. Performance on individual classes

Class	Recall (%)	Precision (%)
0	97.98	98.30
1	99.59	93.37
2	96.46	97.81
3	94.58	96.93
4	96.99	96.35
5	97.87	95.57
6	97.81	98.24
7	95.82	96.26
8	94.92	96.98
9	93.89	96.55

Table 6. The performance comparison of the well-known handwritten digit recognition systems with the proposed method

Rank	Method	Precision (%)	Normalized digits
1	Proposed method	96.64	No
2	Gattal *et al*. 2014 [7]	96.62	No
3	Multi-Scale oBIF column [19]	95.84	No
4	LPQ and LBP features [12]	95.36	No
5	oBIFs and background features [9]	95.21	No
6	Jadavpur [6]	94.75	No
7	Tébessa I [6]	77.53	No

The experimental results unequivocally indicate the superiority of the LDNP column scheme in handwritten digit recognition. By effectively combining global and localized feature extraction methods and employing a strategic combination of features, our approach not only achieves high recall rates but also maintains consistent performance across varied digit classes. The comparative analysis further underscores the method's relevance and potential in advancing the area of digit recognition.

5. CONCLUSION

The goal of this study is to enhance the recognition rates of isolated digit recognition systems by enhancing feature extraction. The focus of the investigation was the proposed Column scheme-based LDNP, which utilizes derivative Gaussian and Kirsch masks. Features were extracted from various parameter settings, including the WI and different image regions, using UGS. Classification was then conducted using SVM. The results show promising outcomes, indicating the effectiveness of the LDNP column scheme in improving recognition accuracy. Our research explored various parameter settings and feature extraction methods, with SVM classification serving as a robust framework. These findings not only validate the potential of the LDNP column scheme, but also pave the way for further exploration in digit recognition. In the future, we plan to delve into a wider range of features and implement sophisticated feature selection mechanisms. This strategic approach aims to identify the most impactful features and push the boundaries of recognition accuracy. As we reflect on the complexities encountered in this study, the LDNP column scheme emerges as an innovative beacon in pattern recognition methodologies. Moving forward, the insights gained from this research will guide the development of more accurate, efficient, and adaptable digit recognition systems.

REFERENCES

[1] A. Dundar, J. Jin, and E. Culurciello, "Convolutional clustering for unsupervised learning," *arXiv*, 2015, doi: 10.48550/arXiv.1511.06241.
[2] J. Denker, H. Drucker, I. Guyon, and B. Laboratories, "Comparison of learning algorithms for handwritten digit recognition," in *International Conference on Artificial Neural Networks*, 1995, pp. 53–60.
[3] Y. Yamashita and T. Wakahara, "Affine-transformation and 2D-projection invariant k-NN classification of handwritten characters via a new matching measure," *Pattern Recognition*, vol. 52, pp. 459–470, 2016, doi: 10.1016/j.patcog.2015.10.002.

[4] F. Lauer, C. Y. Suen, and G. Bloch, “A trainable feature extractor for handwritten digit recognition,” *Pattern Recognition*, vol. 40, no. 6, pp. 1816–1824, 2007, doi: 10.1016/j.patcog.2006.10.011.

[5] A. G. Hochuli, L. S. Oliveira, A. S. Britto Jr, and R. Sabourin, “Handwritten digit segmentation: Is it still necessary?, ” *Pattern Recognition*, vol. 78, pp. 1–11, 2018, doi: 10.1016/j.patcog.2018.01.004.

[6] M. Diem, S. Fiel, A. Garz, M. Keglevic, F. Kleber, and R. Sablatnig, “ICDAR 2013 competition on handwritten digit recognition (HDRC 2013),” in *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR*, 2013, pp. 1422–1427, doi: 10.1109/ICDAR.2013.287.

[7] A. Gattal, Y. Chibani, C. Djeddi, and I. Siddiqi, “Improving isolated digit recognition using a combination of multiple features,” in *Proceedings of International Conference on Frontiers in Handwriting Recognition, ICFHR*, 2014, pp. 446–451, doi: 10.1109/ICFHR.2014.81.

[8] S. Aly and A. Mohamed, “Unknown-Length Handwritten Numeral String Recognition Using Cascade of PCA-SVMNet Classifiers,” in *IEEE Access*, vol. 7, pp. 52024-52034, 2019, doi: 10.1109/ACCESS.2019.2911851.

[9] A. Gattal, C. Djeddi, Y. Chibani, and I. Siddiqi, “Isolated handwritten digit recognition using oBIFs and background features,” in *Proceedings-12th IAPR International Workshop on Document Analysis Systems, DAS 2016*, 2016, pp. 305–310, doi: 10.1109/DAS.2016.10.

[10] N. Ilmi, W. T. A. Budi and R. K. Nur, “Handwriting digit recognition using local binary pattern variance and K-Nearest Neighbor classification,” 2016 *4th International Conference on Information and Communication Technology (ICoICT)*, Bandung, Indonesia, 2016, pp. 1-5, doi: 10.1109/ICoICT.2016.7571937.

[11] L. Heutte, T. Paquet, J. V. Moreau, Y. Lecourtier, and C. Olivier, “A structural/statistical feature based vector for handwritten character recognition,” *Pattern Recognition Letters*, vol. 19, no. 7, pp. 629–641, 1998, doi: 10.1016/S0167-8655(98)00039-7.

[12] A. Gattal and F. Abbas, “Isolated handwritten digit recognition using LPQ and LBP features,” in *ACM International Conference Proceeding Series*, 2020, doi: 10.1145/3447568.3448465.

[13] L. D. Griffin and M. Lillholm, “Symmetry sensitivities of derivative-of-gaussian filters,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 6, pp. 1072–1083, 2010, doi: 10.1109/TPAMI.2009.91.

[14] V. Ojansivu and J. Heikkilä, “Blur insensitive texture classification using local phase quantization,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, pp. 236–243, 2008, doi: 10.1007/978-3-540-69905-7_27.

[15] T. Ojala, M. Pietikäinen, and T. Mäenpää, “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, doi: 10.1109/TPAMI.2002.1017623.

[16] B. Gudla, S. R. Chalamala, and S. K. Jami, “Local binary patterns for gender classification,” in *Proceedings-AIMS 2015, 3rd International Conference on Artificial Intelligence, Modelling and Simulation*, 2016, pp. 19–22, doi: 10.1109/AIMS.2015.13.

[17] E. Al-wajih and R. Ghazali, “An enhanced LBP-based technique with various size of sliding window approach for handwritten Arabic digit recognition,” *Multimedia Tools and Applications*, vol. 80, pp. 24399–24418, 2021, doi: 10.1007/s11042-021-10762-x.

[18] A. J. Newell and L. D. Griffin, “Writer identification using oriented basic image features and the delta encoding,” *Pattern Recognition*, vol. 47, no. 6, pp. 2255–2265, 2014, doi: 10.1016/j.patcog.2013.11.029.

[19] A. Gattal, C. Djeddi, A. Jamil, and A. Bensefia, “Multi-scale oriented basic image features column for handwritten digit recognition,” *Advances in Intelligent Systems and Computing*, vol. 1383 AISC, pp. 289–298, 2021, doi: 10.1007/978-3-030-73689-7_28.

[20] A. R. Rivera, S. Member, J. R. Castillo, and S. Member, “Local directional number pattern for face analysis: face and expression recognition,” *Ieee Transactions on Image Processing*, pp. 1–13, 2011.

[21] I. Bendib, A. Gattal, and G. Marouane, “Handwritten digit recognition using deep CNN,” in *ACM International Conference Proceeding Series*, 2020, doi: 10.1145/3432867.3432896.

[22] S. Rao N and N. K. Babu C, “Enhanced ResNet-151-based fused features for optimized Bi-LSTM-DNN-aided handwritten character and digits recognition,” *Expert Systems with Applications*, vol. 244, 2024, doi: 10.1016/j.eswa.2023.122860.

[23] V. Agrawal, J. Jagtap, S. Patil, and K. Kotecha, “Performance analysis of hybrid deep learning framework using a vision transformer and convolutional neural network for handwritten digit recognition,” *MethodsX*, vol. 12, 2024, doi: 10.1016/j.mex.2024.102554.

[24] C. L. Liu, K. Nakashima, H. Sako, and H. Fujisawa, “Handwritten digit recognition: Benchmarking of state-of-the-art techniques,” *Pattern Recognition*, vol. 36, no. 10, pp. 2271–2285, 2003, doi: 10.1016/S0031-3203(03)00085-2.

[25] Z. Abdelmoumene, L. Lakhdar, and G. Abdeldjamil, “Handwritten digit recognition using encryption methods,” in *4th International Conference on Pattern Analysis and Intelligent Systems, PAIS 2022-Proceedings*, 2022, doi: 10.1109/PAIS56586.2022.9946886.

[26] T. Jabid, M. H. Kabir, and O. Chae, “Local directional pattern (LDP) for face recognition,” *International Journal of Innovative Computing, Information and Control*, vol. 8, no. 4, pp. 2423–2437, 2012.

[27] B. Zhang, Y. Gao, S. Zhao, and J. Liu, “Local derivative pattern versus local binary pattern: face recognition with high-order local pattern descriptor,” *IEEE Transactions on Image Processing*, vol. 19, no. 2, pp. 533–544, 2010.

[28] R. A. Kirsch, “Computer determination of the constituent structure of biological images,” *Computers and Biomedical Research*, vol. 4, no. 3, pp. 315–328, 1971.

[29] J. Sauvola and M. Pietikäinen, “Adaptive document image binarization,” *Pattern Recognition*, vol. 33, no. 2, pp. 225–236, 2000, doi: 10.1016/S0031-3203(99)00055-2.

[30] J. T. Favata and G. Srikantan, “A multiple feature/resolution approach to handprinted digit and character recognition,” *International Journal of Imaging Systems and Technology*, vol. 7, no. 4, pp. 304–311, 1996, doi: 10.1002/(SICI)1098-1098(199624)7:4<304::AID-IMA5>3.0.CO;2-C.

[31] V. N. Vapnik, “The nature of statistical learning theory,” *The Nature of Statistical Learning Theory*, 2000, doi: 10.1007/978-1-4757-3264-1.

BIOGRAPHIES OF AUTHORS

Mohammed Aouine was born in Algeria. Received the Engineer degree in Computer Science from University of Annaba (Algeria) in 2004, M.S. degree in Computer Science "IA and imagery" from University of Guelma (Algeria) in 2009. Currently, he is working as associate professor at the Department of Mathematics and Computer Science in University of Tebessa (Algeria). He supervised many Master and License students. In addition, he has collaborated as a member on several research projects. His research interests include image analysis, pattern recognition, and recognition of handwriting. He can be contacted at email: mohamed.aouine@univ-tebessa.dz.

Abdeljalil Gattal was born in Algeria. He received his B.S. degree in Computer Science from University of Skikda (Algeria) in 2004, M.S. degree in Computer Science Information and Knowledge Systems" from Abbes Laghrour University of Khenchela (Algeria) in 2009 and he received his Ph.D. in 2016 from Ecole Nationale Supérieure d'Informatique (ESI-Algeria) in Computer Science and focuses in segmentation-verification for handwritten digit recognition. Currently, he is working as full professor at the Department of Mathematics and Computer Science in University of Tebessa (Algeria). Currently, he is a leader of the Laboratoire de Vision et d'Intelligence Artificielle (LAVIA) at the University of Tebessa. He supervised many master and license students. He has published a number of papers. In addition, he has collaborated as a member on several research projects and also participated in several scientific competitions. His research interests include image analysis, pattern recognition, and recognition of handwriting. He can be contacted at email: abdeljalil.gattal@univ-tebessa.dz.

Chawki Djeddi was born in Algeria. He is presently working as full professor in the Department of Mathematics and Computer Science, University of Tebessa, Tebessa, Algeria. He received his Ph.D. in 2014 from Badji Mokhtar-Annaba University, Annaba, Algeria and specializes in document image analysis and recognition. His research interests include image processing and pattern recognition with applications to document image analysis, content-based image retrieval and signature verification on disguised signatures and skilled forgeries. He has been regularly participating in the top conferences in areas of document analysis and handwriting recognition. Currently, he is a research team leader of the Laboratoire de Vision et d'Intelligence Artificielle (LAVIA) at the University of Tebessa. He has also supervised a number of masters theses. As a part of his professional activities, in addition to teaching, he also takes up several administrative responsibilities as requested and when needed. He can be contacted at email: c.djeddi@univ-tebessa.dz.

Faycel Abbas was born in Algeria. Received a Ph.D. in Computer Science from Akli Mohand Oulhadj University, Bouira, Algeria in 2022. Following his Ph.D., he worked as an assistant professor in the Department of Mathematics and Computer Science, University of Tebessa, Tebessa, Algeria, where he did research in the areas of biometrics, writer identification, image processing, and document analysis. He has published various papers in the above general areas. He can be contacted at email: Faycel.Abbas@univ-tebessa.dz.