A Novel Method for Banknote Recognition Using a Combined Histogram of Oriented Gradients and Scale-Invariant Feature Transform

Gawaher S. Hussein
Information System Department, Faculty of Computers and Informatics, Zagazig University, 44523, Zagazig, Egypt, dralwathiq2012@gmail.com

Sherif Elseuofi
Information System Department, Higher Institute for Computers & Specific Studies, Ras Elbar, Damietta, 34737, Egypt, dralwathiq2012@gmail.com

Wathiq H. Dukhan
Mathematic Department, Faculty of Science, Sanaa University, Sanaa, Yemen, dralwathiq2012@gmail.com

Asmaa H. Ali
Information System Department, Faculty of Computers and Informatics, Zagazig University, 44523, Zagazig, Egypt, dralwathiq2012@gmail.com

Follow this and additional works at: https://digitalcommons.aaru.edu.jo/isl

Recommended Citation
S. Hussein, Gawaher; Elseuofi, Sherif; H. Dukhan, Wathiq; and H. Ali, Asmaa (2023) "A Novel Method for Banknote Recognition Using a Combined Histogram of Oriented Gradients and Scale-Invariant Feature Transform," Information Sciences Letters: Vol. 12 : Iss. 9 , PP -. Available at: https://digitalcommons.aaru.edu.jo/isl/vol12/iss9/19

This Article is brought to you for free and open access by Arab Journals Platform. It has been accepted for inclusion in Information Sciences Letters by an authorized editor. The journal is hosted on Digital Commons, an Elsevier platform. For more information, please contact rakan@aaru.edu.jo, marah@aaru.edu.jo, u.murad@aaru.edu.jo.
A Novel Method for Banknote Recognition Using a Combined Histogram of Oriented Gradients and Scale-Invariant Feature Transform

Gawaher S. Hussein¹, Sherif Elseoufi², Wathiq H. Dukhan³*, and Asmaa H. Ali³

¹Information System Department, Faculty of Computers and Informatics, Zagazig University, 44523, Zagazig, Egypt
²Information System Department, Higher Institute for Computers & Specific Studies, Ras Elbar, Damietta, 34737, Egypt
³Mathematic Department, Faculty of Science, Sana’a University, Sana’a, Yemen

Received: 14 Mar. 2023, Revised: 22 Jul. 2023, Accepted: 11 Aug. 2023. Published online: 1 Sept. 2023.

Abstract: Automated banknote recognition systems are essential for people with visual impairments who face challenges distinguishing between different currency denominations. This study presents a novel method aimed at helping blind people identify banknotes from three different countries (Egypt, Saudi Arabia, and the United States of America) by using a proposed feature detection algorithm. Our proposed system has two main stages: the proposed algorithm uses the Speeded-UP Robust Features (SURF) algorithm for key point detection, as it is fast and robust to variations in geometry and lighting. Then, it extracts features using the scale-invariant feature transform (SIFT) and histogram of oriented gradients (HOG) algorithms, which are scale invariant. This algorithm aims to overcome the limitations of both the SURF and SIFT algorithms and reduce the average response time and computational cost of the SIFT and HOG algorithms. We developed a banknote dataset with 12 classes for three countries. The accuracy of the banknote recognition was 99.2%. The performance of the proposed dataset was compared with that of the global Kaggle Egyptian dataset, resulting in 98.9% accuracy.

Keywords: SIFT, SURF, HOG, Feature detector, Egyptian Kaggle dataset, Support vector machine (SVM).

1 Introduction

Physical money, such as coins and bills, is used in almost all countries. Banknote recognition is critical for applications such as helping blind people identify their money, and counting machines and ATMs to classify what is being deposited. This research aimed to determine whether banknote recognition can be achieved using methods such as histogram of oriented gradients (HOG) and scale-invariant feature transform (SIFT) algorithms. The SIFT algorithm is a feature-selection method that identifies certain points of an object that remain unchanged when the image is scaled or rotated. This technique is robust to illumination changes, noise, and minor shifts from one viewpoint [1]. There are approximately 285 million visually impaired people worldwide, 39 million of whom are blind. A total of 5.9 million blind individuals resides in Africa, 3.2 million in the United States, and 2 million in Europe [2]. Sight is essential to human survival. Vision is not just a skill that facilitates activities but also affects human behavior. Blindness affects psychological behaviors; a blind person is more likely to experience sadness than a person with limited or normal vision. In addition, they are more likely to experience worry [3] and lack of social contact [4]. Determining the value of a banknote is a challenging task for blind individuals. Therefore, we designed an innovative system based on computer vision and camera technology that automatically identifies banknotes for the visually impaired. This paper presents an innovative approach to identifying currencies, such as Egyptian pounds, Saudi riyals, and U.S. dollars, even when they are partially folded, worn, wrinkled, or visible because of their frequent use. Recognizing banknotes in a variable environment is challenging because uncontrolled conditions can affect image quality. Our method addresses this difficulty by accurately recognizing currencies under various conditions. The main contributions to this study are as follows:

• A new banknote recognition method supported by the SIFTS and HOG features was combined.
• Banknotes are handled in several cases such as rotation, illumination, scaling, cluttered backgrounds, and occlusion.
• The proposed algorithm achieves a high true recognition rate of 99.2%. Blind people also tested our banknote recognition system.

*Corresponding author e-mail: dralwathiq2012@gmail.com
A new proposal dataset for banknotes consists of 12 classes: Egyptian pounds (5 EGP, 10 EGP, 20 EGP, and 100 EGP), Saudi riyals (5 Saudi riyals, 10 Saudi riyals, and 100 Saudi riyals), and US dollars (5 dollars, $10, $20, $50, and $100). Images of banknotes were captured under various settings and lighting conditions. The dataset consisted of both new and old notes.

The proposed system was evaluated using two datasets: one collected by the authors and the other from the Global Kaggle currency database.

The remainder of this paper is organized as follows: Section 2 reviews the relevant literature. The proposed system is discussed in Section 3. The experimental setup and results are presented in Section 4. Finally, Section 5 concludes the paper.

2 Literature Review

Many researchers have contributed to the development of methods for recognizing currency. The different characteristics of coins and bills require different approaches for recognition, and this section provides an overview of existing currency recognition techniques. The authors of [5] proposed an image-processing system for currency recognition using four algorithms (SIFT, FAST, ORB, and SURF). Following the discussion, it was discovered that all algorithms have advantages and disadvantages. Instead of selecting the most suitable algorithm for image processing, these algorithms can be integrated and applied to obtain optimal results. The experimental analysis comparison based on rotation, illumination, and scale invariance indicated that SIFT and SURS are the two best techniques. The authors of [6] created a component-based system based on SURF. They used a dataset of 140 images to mimic the real-world application scenario, which included 20 images for each dollar currency denomination (1, 5, 10, 50, and 100) obtained from various situations. The system achieved 100% accuracy, but the implementation time was long, and their dataset included only U.S. banknotes.

The authors of [7] proposed a system for detecting paper currency. This method combines LBP and SURF features. The system was trained to identify the banknotes. A trained SVM classifier was used for the prediction. The method works only for the Bangladeshi taka (Bangladeshi currency). In addition to identifying currencies, this system can compute the total amount of cash in an image (when the image has more than one currency). This method has an accuracy rate of 92.6%; however, there are other techniques. The authors of [8] proposed a new technique for feature detection to enhance the performance of these methods. They combined SITF and SURF algorithms to make the proposed method faster and more robust. According to the results, better results were achieved (with higher speed and robustness) than those of the SIFT and SURF algorithms. The authors of [9] introduced the SURF descriptor method coupled with an improved FAST feature point. In the experiments, the proposed method was compared with the SURF and SIFTS algorithms. Similarly, the proposed solution extracted more feature points than the FAST method. The proposed solution was 100 times faster than the traditional SIFT method. The recognition speed was 10 times higher than that of the SURF method. This approach was successful in achieving faster matching and higher accuracy than the SIFT algorithm. The HOG feature was extracted from the watermark by the system using latent images and microprinting and classified using an SVM algorithm, as suggested by [10]. The system achieved 100% accuracy. However, it was tested on 500 and 1,000 BDT bills.

The authors suggested a method for recognizing Nigerian paper currencies based on an SVM and genetic algorithms [11]. The GA optimizes the parameters, and the SVM-based classifier classifies the currency. This method does not test other global currencies and lacks a good feature set for classification. The authors presented a mobile phone-based system for identifying paper currencies as an aid for people with impaired vision [12]. The method uses k-means clustering to identify the U.S. The system extracts computationally intensive SIFT features from the note, thereby boosting the recognition time. Similarly, [13] created a mobile phone method for identifying Indian banknotes. The recognition technique uses the GrabCut algorithm and the visual BOW recognition method to separate the foreground from the background. The overall accuracy was 96.7% when 2584 images were used.

Moreover, the approach was subject to failure due to image lighting and background inclusion in the captured banknote image. In addition, [14] developed a currency recognition system. They used SIFT features from color and grayscale images and compared their recognition rates. They tested their method by using 400 Jordanian cash denominations (coins and banknotes). Their method has problems with illumination, depends on the distance of image capture, and is relatively slow on mobile devices. We proposed an innovative system for paper currency detection based on the SURF detector with SIFT and HOG features, using SVM for classification.

3 Methodology

The proposed system was used to identify different currencies to help the blind by using a camera to capture a picture of
the currency through it, and then more than one algorithm was combined to recognize the image of the currency. After
banknote image acquisition, the following was applied first: key point detection (an interesting point in the currency
image) using the SURF algorithm; then, both SIFT and HOG were combined for key point feature extraction and
description. Finally, the SVM algorithm was used for the classification.

3.1 Data collection and description

The camera used to capture the dataset images had a resolution of 448 × 336 pixels. The images varied in terms of
angle, distance, scale, illumination, and distortion (folded and wrinkled). All images were taken from both the front and
back and saved as JPEG, which represented the Joint Photographic Experts Group. Figure 1 shows some of the dataset
images.

3.2 Overall structure

The proposed system was implemented on a database containing 1700 banknote images. The dataset was divided into
training (70%) and testing (30%). The steps of the proposed approach are illustrated in Figure 2.

A. Banknote image acquisition

The workflow sequence always begins with the first step of image acquisition. Therefore, processing cannot be
performed without the use of images. This is the process of creating digital photographs, which most commonly begins
with a real scene. The images used in this study were banknotes captured using a digital camera. The images were then
saved for further preprocessing.

Voice feedback

Recognition result

Banknote image acquisition

Pre-processing

Key point detection (SURF)

Feature extraction (SIFT+HOG)

Classification(SVM)

Voice feedback

Recognition result

Banknote image acquisition

Pre-processing

Key point detection (SURF)

Feature extraction (SIFT+HOG)

Classification(SVM)

Fig. 1: Some of the banknotes of the dataset

Fig. 2: Banknote Recognition System Overview

B. Image pre-processing

- A preprocessing phase was conducted to make the identification of descriptive characteristics more accurate and to
guarantee that the system was capable of accurate recognition. Images from both the training and testing samples
were reduced to less than 20 KB to reduce the amount of time required for processing and to maximize the use of
available RAM. The primary purpose of image preprocessing was to eliminate unwanted distortions and improve
certain image characteristics required for further image processing or analysis. The efficiency of the recognition
system was significantly improved.

- Image adjustment: However, the camera’s image size was insufficient. Image adjustment could be accomplished
using image interpolation. This method can be applied to a variety of functions, including zooming, rotating,
shrinking, and making geometric adjustments.

- Image smoothing: There was some noise in the images after they were acquired through a camera and then
transferred using various methods. Noise elimination is one of the most significant tasks in image preprocessing.
This noise may interfere with pattern matching and segmentation.
C. Key point detection

In this study, the SURF algorithm was used to detect key points, which are distinctive image features that can be detected repeatedly despite changes in the image scale (resolution), illumination, noise, image orientation, and perspective. The main benefit of the SURF algorithm is its rapid interest-point recognition capability. It is also invariant to common image transformations such as rotation, scale, lighting, and tiny perspective changes. The SURF algorithm assigns orientations and scales to each key point using the SURF method [15]. Then, it computes a descriptor that characterizes the neighborhood of the key point. Next, it measures the distance between descriptors that are not located at key points [16]. The SURF algorithm is a faster version of the SIFT algorithm, with similar properties. However, the SURF algorithm is protected by patents.

D. Feature extraction

For banknote feature extraction and description, the SIFT and HOG techniques were used.

• Scale-invariant feature transform

The SIFT algorithm begins by creating a scale space, which is an internal representation of the original image that ensures scale invariance [17]. Next, the Laplacian of the Gaussian operator was applied, which detected corners and edges, to find key points by comparing each pixel with its neighboring pixels in the same image and adjacent images. Because the key points were not exactly located on the pixels, their subpixel locations were computed [18]. Some of the key points were along an edge and not useful; therefore, they were removed. The orientations and scales were then assigned to each key point using gradients. Finally, descriptors were computed for each key point, which facilitated the matching and comparison of the key points [17].

• Histogram of oriented gradient (HOG)

The HOG feature extraction algorithm detects objects in images using histograms of the orientations of the gradients of the local regions of the input brain MR images [19]. The main idea of this descriptor is that the local shape of an object can be characterized by the distribution of local gradients or edge directions without knowing their exact locations in the image [20]. HOG descriptors are key features that encode object characteristics into a sequence of specific numbers that can differentiate objects from each other [21]. The HOG algorithm splits an image into small regions, called cells, and computes a histogram of the gradient orientations for each cell. It then groups several adjacent cells into a block and normalizes the histograms of the cells in the block. The concatenation of all histograms forms a histogram of the oriented gradient (HOG) feature vector of the image [22]. The following steps are involved in extracting the HOG feature vector from an image:

The horizontal \( I_x(i,j) \) and vertical \( I_y(i,j) \) gradients of an image were computed using a gradient filter \([1; 0; 1]\). The gradient magnitude \( I(i,j) \) and angle \( \theta(i,j) \) were then calculated from the horizontal and vertical components.

\[
|I(i,j)| = \sqrt{I_x(i,j)^2 + I_y(i,j)^2} \tag{1}
\]

\[
\theta(i,j) = \arctan\left(\frac{I_y(i,j)}{I_x(i,j)}\right) \tag{2}
\]

The image was divided into \( 8 \times 8 \) pixel cells, and a histogram of nine orientation bins ranging from 0° to 180° was computed for each cell. The bin values are the sum of the magnitudes \( |I(i,j)| \) of the gradients that have the same angle \( \theta(i,j) \) as the bin.

The cells were then grouped into blocks, and the histograms of the cells in each block were normalized using the L2-Hys (Lowe-style clipped L2 norm) method. The HOG feature vector of the image is a concatenation of all the histograms.

E. Combined feature

We concatenated the SURF and HOG features into a single feature vector, which enhanced feature representation.

\[
F_{\text{Combined}} = [F_{\text{SIFT}}, F_{\text{HOG}}] \tag{3}
\]

F. Interest point matching using support vector machine

Following the detection of robust key points using the preceding procedure, the next step was to match the interest points using the SVM. It is a popular classifier scheme applied to classification problems and is based on supervised learning techniques [23, 24]. Several experimental studies have shown that SVMs achieve significantly better results with higher search accuracies [25, 26]. In this study, we used an SVM with a linear kernel because of its simplicity and
computational efficiency for training and classification. The kernel function transforms nonlinear samples into a high-dimensional feature space where they can be separated by a linear hyperplane [27]. The SVM classifier determines the optimal hyperplane that maximizes the margin between the two classes, as shown in Equation 4.

\[
K(x, y) = x^T y + c
\]  

where \( x \) and \( c \) are the hyperplane parameters.

To handle datasets that are not linearly separable, a classifier that can capture non-linear patterns was used. The nonlinear SVM was formulated as follows:

\[
K(x, y) = 1 - 2 \sum_{i=1}^{n} \frac{(x_i - y_i)^2}{x_i + y_i}
\]

Where \( x_i \) denotes the training features and \( y_i \) the label of \( x_i \). The sum is calculated over a subset of \( x_i \). These feature vectors are known as support vectors because they are the closest to the separating hyperplane. Figure 3 shows the point matching using the SVM.

![Fig. 3: Key-point Matching using SVM](image)

G. Audio output display

Every class in the database was linked to an audio file that pronounced the name of the currency, meaning that the display process of the speech file depended on the banknote identification result. Users could customize speech settings such as speed, volume, and language to suit their needs.

4 Results and Discussion

Section 4.1 evaluates the performance of the proposed system under different conditions, such as intensity, rotation, scale, and noise. This section covers four phases. The second part compares the proposed system algorithm with the SURF and the SIFT algorithms in Section 4.2. The third part discusses the evaluation of the proposed system in Section 4.3. Finally, a comparison with other datasets is presented in Section 4.4.

4.1 Towards an evaluation of the proposed system's efficiency

In this section, the sensitivity of the proposed system to different intensities, rotations, scaling, and noise are discussed.

- **Intensity**
  In this section, banknote images are matched with varying intensities using the proposed system. Table 1 displays the results of banknote images with varying intensities, as shown in Figure 4.

- **Rotation**
  In this section, banknote images are matched with a 45-degree rotation using the proposed system. Table 2 presents the matching results. Figure 5 shows the matching of the original image with its rotation.

- **Scaling**
  In this section, banknote images were matched at different scales to examine the impact of matching on scaling using the proposed system (Table 3; Figure 6).
• Noisy images

In this section, 40% noise was added to the original image to evaluate the effect of noise on the matching rate. The results are presented in Table 4 and Figure 7.

![Image matching with varying intensities using the proposed system.](image1)

**Fig. 4:** Image matching with varying intensities using the proposed system.

<table>
<thead>
<tr>
<th></th>
<th>Time (Sec)</th>
<th>K-Points 1</th>
<th>K-Points 2</th>
<th>Matches</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed System</td>
<td>0.04</td>
<td>166</td>
<td>160</td>
<td>159</td>
<td>99.4</td>
</tr>
</tbody>
</table>

![Image matching for the original and scaled images using the proposed system.](image2)

**Fig. 6:** Image matching for the original and scaled images using the proposed system.

![Matching the original image towards its rotated image using proposed system](image3)

**Fig. 5:** Matching the original image towards its rotated image using proposed system.

<table>
<thead>
<tr>
<th></th>
<th>Time (Sec)</th>
<th>K-Points 1</th>
<th>K-Points 2</th>
<th>Matches</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed System</td>
<td>0.02</td>
<td>170</td>
<td>150</td>
<td>148</td>
<td>98.6</td>
</tr>
</tbody>
</table>
Table 3: The results of comparing the image to its scaled counterpart.

<table>
<thead>
<tr>
<th>Proposed System</th>
<th>Time (Sec)</th>
<th>K-Points 1</th>
<th>K-Points 2</th>
<th>Matches</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.10</td>
<td>120</td>
<td>100</td>
<td>98</td>
<td>98</td>
</tr>
</tbody>
</table>

4.2 Compare the proposed system algorithm

In this section, the proposed system is compared with both SIFT and SURF algorithms, as shown in Table 5. The results showed that SIFT is better than SURF for key point detection, but it takes longer, whereas the SURF method takes less time. The resulting technique was improved by combining these two approaches.

Table 4: The results of image matching with 40% noise added.

<table>
<thead>
<tr>
<th>Proposed System</th>
<th>Time (Sec)</th>
<th>K-Points 1</th>
<th>K-Points 2</th>
<th>Matches</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.05</td>
<td>350</td>
<td>320</td>
<td>319</td>
<td>99.7</td>
</tr>
</tbody>
</table>

4.3 Evaluation of the proposed system

The proposed system was implemented in a database containing 12 classes of banknotes belonging to three different countries, with each class consisting of 100 images of the same banknote. The total number of images in the database was 1700. A total of 1190 images were used in the training process, and 510 images were used for testing. To evaluate the performance of the proposed method, a training dataset was used, and 506 images were successfully identified from 510 images. The accuracy of the proposed system was calculated using the recognition rate, as follows:

\[
\text{Recognition rate} = \frac{\text{Number of correctly recognized items}}{\text{Total number of items}} \times 100\%.
\]

4.4 Comparison with other datasets to affirm efficiency of the proposed system

The proposed system was applied to the most famous Egyptian currency dataset, and the results obtained proved the accuracy of the system. Table 6 shows the original Kaggle image, and the image obtained after applying a set of
changes. Table 7 shows the performance evaluation of the proposed method on intensities, rotations, scaling, and noise using the Egyptian Kaggle dataset. Figure 8 illustrates the experience of using the Egyptian Kaggle database with the proposed system.

| Table 6: Some transformations on an original image from the Kaggle database |
|---|---|
| (a) The original Kaggle image | (b) 90-degree rotation image |
| (c) Image at various scales | (d) Image blur |
| (e) Image with many hues | (f) Image with varying saturation and value |
| (g) Image warping | (h) Image of RGB noise |

5 Conclusions and Future Work

The current recognition research is expanding daily, and various algorithms have been employed to achieve more accurate results. The current study suggests a new banknote recognition technique supported by a SURF detector with SIFT and Hog features. A support vector machine (SVM) was suggested as the classifier in this study. The proposed banknote recognition system was evaluated by 20 blind participants. Experiments on a combined banknote image database containing 12 denominations from three different countries yielded a classification accuracy of 99.2%, demonstrating that the proposed system outperforms previously used methods. To evaluate the performance of the proposed approach, we applied it to the Kaggle dataset of Egyptian banknotes and obtained a remarkable accuracy of
98.9%. As part of our future work, we intend to explore the integration of SURF features with other techniques, such as color histograms, different clustering methods, such as PCA and decision trees, and deep learning using CNN, which might further enhance accuracy. Additionally, we will use meta-heuristic algorithms to recognize banknotes that are incomplete and examine the impact of such methods on automated banknote recognition [28-30].

Table 7: Performance evaluation of the proposed method using the Egyptian Kaggle dataset

<table>
<thead>
<tr>
<th>Image</th>
<th>Proposed method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Image a feature extraction</td>
</tr>
<tr>
<td>Image (a) and (b)</td>
<td>878</td>
</tr>
<tr>
<td>Image (a) and (c)</td>
<td>822</td>
</tr>
<tr>
<td>Image (a) and (d)</td>
<td>878</td>
</tr>
<tr>
<td>Image (a) and (e)</td>
<td>878</td>
</tr>
<tr>
<td>Image (a) and (f)</td>
<td>878</td>
</tr>
<tr>
<td>Image (a) and (g)</td>
<td>712</td>
</tr>
<tr>
<td>Image (a) and (h)</td>
<td>871</td>
</tr>
</tbody>
</table>

Fig. 8: The experience using the Egyptian Kaggle database.


Conflict of interest

The authors declare that there is no conflict regarding the publication of this paper.

References


