International Journal of Mathematics and
Computer Science, **16**(2020), no. 2, 577–588

Metaheuristic algorithm for Capital Letters Images Recognition

Muna Jaffer Al-Shamdeen, Ansam Nazar Younis, Hiba Adreese Younis

Department of Computer Science
University of Mosul
Mosul, Iraq

email: muna.jaffer@uomosul.edu.iq, anyma8@uomosul.edu.iq,
hibaadreese2017@gmail.com

(Received September 7, 2020, Accepted October 22, 2020)

Abstract

This work aims at identifying the pattern of English capital letters printed in JPG graphic images via a computer system that converts these images and edits them electronically to written characters. The system identifies the image of the input character and determines what this character is. It uses for this purpose a smart algorithm (Particle swarm optimization (PSO)) after being developed and improved by adding a method to calculate the fitness of each Individual of the swarm, which helps to accelerate the search for the goal and reach the optimal solution.

1 Introduction

The huge growth in the Internet and computers has led to the development of useful algorithms and software that support this development to help users search within different communication sites and retrieve the required documents [1]. However, most natural languages are complex languages and the size of their data is of enormous proportions [2] and thus the need for

Key words and phrases: Morphological Methods, Dilation, Erosion, Histogram Equalization, PSO, NPSO.
AMS (MOS) Subject Classifications: 68T10.
ISSN 1814-0432, 2021, http://ijmcs.future-in-tech.net

sophisticated algorithms in line with this development and capable of discriminating within a relatively short period of time. The use of intelligent algorithms that simulate the lives of living creatures trying to live in swarms and how they search for shelter or food and act smartly so that they can produce low-cost, fast and powerful solutions to many complex and major problems [3]. The term metaheuristic was introduced to describe a heuristic high level used to lead other heuristics for a better search space evolution [4]. PSO is considered one of the algorithms of metaheuristic [5][6]. An OCR system typically includes few main stages, starting with Primary treatment followed by segmentation and feature extraction ending with classification [7]. Husnain and Naweed have introduced the Bayesian decision theory in classification to recognize the pattern of the English capital alphabet [8]. Dash and Natak proposed an algorithm for hand written English alphabet pattern recognition which is based on the principle of Artificial Neural Network (ANN)[9]. Perwej and Chaturvedi proposed an English alphabet recognition system using neural networks [10]. Adamu and et al. designed a program for English characters recognition using multilayer perceptron neural network with a single hidden layer [11]. The aim of the research is the development of an intelligent system to facilitate the computer's task for recognizing English letters in an efficient and fast manner as an important step in the path of interaction between natural language and machine. It is regarded as an important step in the field of interpretation of complex natural languages.

2 Morphological Methods

Binary images can contain uncounted faults. Morphology is a wide extent of operations that modifies the images based on shapes. It can be considered as one of data processing methods used in image processing. Morphological methods check the image with a small set called structuring element, which is used to all probable input image positions and produce the same output size [12]. This operation generates a new binary image that will contain a non-zero pixel value in the similar position in the input image. There are various structuring element such as square shaped, cross and diamond shaped. The main morphological operations are dilation, erosion, opening, and closing [13]. Binary Dilation is one of mathematical morphology operators that can be usually implemented on binary image, and there are other versions that can be implemented on gray scale images. The primary influence for this operator is to expand the desired pixel of area boundaries in which these pixels increase in size whereas holes within these area be smaller. In dilation, the white pixel in the image augmented which makes it appears broader. All background pixels which are touching an object pixel is transformed into an object pixel [14]. Erosion is another mathematical morphology operators that can be usually implemented on binary image, and there are other versions that can be implemented on gray scale images. The basic influence for it on a binary image is to erode away the boundaries of areas of foreground pixels, in which areas of foreground pixels shorten in size, and holes within those areas become wider [14]. Morphological Filtering is performed by considering opening and closing operations as filters. These operations can filtrate every detail whose size is smaller than the size of structuring element of an image, opening can filter the binary image based on the size of the structuring element in which image portions that are suitable for the structuring element are exceeded by the filter whereas smaller structures are blocked up and eliminated from the produced image. The structuring element size plays an important role in getting off noisy and unwanted details without destroying the image interested objects [15], [16].

3 Histogram Equalization

This is one type of image enhancement technique used to improve the contrast of digital images. It transforms the input images in such a way that the output images have equalized histogram and an approximately uniform distribution of pixel values [17]. In this approach, we first find the occurrences for each intensity in the image. We then find the probability density function (PDF) by dividing occurrences for each intensity by the total number of pixels. Afterwards, we find the cumulative density function (CDF) whose first value is equal to the first value of PDF and its other value is found by adding the CDF in the previous step with the PDF in the present step. Finally, we multiply each CDF with the maximum gray value to produce new gray levels and assign new values for each gray value in the image [18], [19].

4 Traditional PSO Algorithm

This is an intelligent algorithm for evolutionary improvement inspired by the nature of living creatures widely used to solve complex problems to reach the target especially with non-linear functions and optimized way [20]. This algorithm enables the search in very large areas and gives very little or no assumptions about the optimal solution. This algorithm is therefore used to solve complex and partially irregular problems that are often noisy [21]. The algorithm treats each solution as a particle and has two important properties: location l, and velocity ve, the two attributes are associated with each particle. Therefore, l=(l_i1,l_i2,...,l_iN) and ve=(ve_i1,ve_i2,...,ve_iN), where N represents the dimension of the problem and at each stage of the search for a solution, a fitness function is given to the particles in the swarm. The speed and location values are updated according to the two equations:

$$v = w.v + c_1 m_1(pBest - 1) + c_2 m_2(nBest - 1) \tag{4.1}$$

$$l = l + \Delta t \tag{4.2}$$

w represents the weight of inertia which is responsible for controlling past velocities effects of particles, c_1 ,c_2, represent positive constants which are called acceleration parameters, m_1 ,m_2, random values take new values in each appearance, Δt represents the time step, $pBest$ is the best current location that has reached or passed the particle until the present time while $nBest$ is the best current location that a particle of neighboring particles has reached or passed until the present time [21], [22].

5 Work Stages

5.1 Data Collection Stage

For written English characters, there are many forms and font styles that belong to the same character. So a special database was created for a group of 26 large and printed English letters (A to Z), which included ten different models of each letter, 70 percent of them were used in the training process and 30 percent in the testing process. A set of strange pictures about this dataset was also selected including 20 non dataset different model, 10 of which are actually true and belong to the same types of uppercase letters within the dataset, and the other 10 are non-dataset models containing strange symbols that do not belong to uppercase classes within the dataset and all these images have been tested in this system.

5.2 Primary Treatment Stage

Primary treatment is a very important part of Optical Character Recognition (OCR) systems in capital letters languages. It is recommended to use morphological methods and histogram equalization to increase the basic image enhancement. After reading the images entered in the system, these images are preprocessed with various stages to obtain clear images that facilitate distinguish them by NPSO algorithm.

5.3 Recognition stage using New Particle Swarm Optimization algorithm (NPSO):

The algorithm has been developed so that the fitness of each member of the swarm is calculated by using one of the Template matching methods which is the method of calculating the amount of correlation between the individual and the target using the correlation coefficient. In the beginning, the first community is entered to be calculated the first generation of it which represents the set of capital English letter images after the Primary treatment that was performed on them, which are located within the database in the system. The best solution is calculated from among the solutions available within the first generation, using the fitness value calculation for each member of society in the first generation, according to the correlation equation between the goal and every member of the first generation which is represented by the following equation from [23]:

$$Corr = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M}(x_{ij}-\overline{x})(y_{ij}-\overline{y})}{\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{M}(x_{ij}-\overline{x})^2}\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{M}(y_{ij}-\overline{y})^2}} \tag{5.3}$$

where $\overline{x}$, $\overline{y}$ represents the mean of image x and y and can be found by the following equation:

$$\overline{x} = \sum_{i=1}^{N}\sum_{j=1}^{M}\frac{x_{ij}}{NM} \tag{5.4}$$

and

$$\overline{y} = \sum_{i=1}^{N}\sum_{j=1}^{M}\frac{y_{ij}}{NM} \tag{5.5}$$

where M, N are the image dimensions. After that, the velocity value is calculated for each member of the current community and thus the first stage of the proposed NPSO algorithm has been completed. The second stage of the algorithm begins by examining the value of the maximum iteration whether the number of the iteration of the algorithm has reached the maximum number or not. If the value has reached the maximum iteration, this means that the algorithm has completed its work and reached the state of stability to the best possible solution. In other cases, the second stage of the algorithm is completed. The new solution is calculated (which means the calculation of the following generations in each cycle) where the velocity value is calculated for each member of the current society and then finding the best possible solution among the expected solutions within the members of society after calculating the fitness value for each individual and finding the best match between the goal and the community members. After that, the new solution is checked to see if the new proposed solution has a higher match value than the old proposed solution or not. If it is actually better than the old solution, then it is chosen as the best current solution instead of the old solution and then the velocity values are updated for all members of the current society and the choice of the improved generation as a new generation which represents the third stage of the algorithm. However, if the current solution is not better than the old solution, then it must return to the beginning of the second stage again. Thus, until the best possible solution is reached which represents the best match between the goal and members of society.

6 Results And Analysis

6.1 Training Stage

The proposed new algorithm was applied on 182 capital letters which represent about 70 percent of all images in the dataset. The results showed the ability of the algorithm to distinguish correctly and with a high capacity, where the percentage of discrimination at this stage was 100 percent.

6.2 Testing Stage

The number of tested images used in the system are 78 different images for various character categories used in the dataset of the system which represents about 30 percent of all images in the dataset. The correct discrimination rate was 97.5 percent, whereas the false discrimination rate was 2.5 percent according to the equations in [24], [25]:

$$\text{Discrimination Rate}(DR) = \frac{\text{Number of samples discriminated correctly}}{\text{Number of all samples}}, \tag{6.6}$$

and

$$\text{False Rate}(FR) = \frac{\text{Number of samples discriminated incorrectly}}{\text{Number of all samples}}. \tag{6.7}$$

Table 1: Results of the strange input images

No.	**Rate(Measure Name)**	**Value(Percent)**
1	DR	97.5
2	FR	2.5
3	Recall	90
4	Spe	90
5	FNA	10
6	Pre	90
7	NPA	90
8	FPA	10

However, DR measure describes the capacity of the system to accurately identify and recognize the shape of the English letter of the images entered into the system that wanted to find its type. So the higher value of this scale means that the system's ability to recognize the English letter in the input image is better. The statistical value of the FR measure represents negative recognition towards the system. Therefore, the lower value of this measure indicates a higher efficiency of the system, where three models have been tested for each type, which comprises 30 percent of the images selected for each type of image used in the database.

Although the DR and FR values shown in the table 1 are the values resulting from the execution of equations 6 and 7 and the calculation of the system's ability to identify 30 percent of the images chosen for the examination process for each type of English letter, such that each type is measured separately. The threshold amount used in the system to distinguish between correct capital letter images and false capital letter images was 0.3. Twenty strange images were also examined by the system to ensure the ability of it to find images of the items in the dataset, the results are shown in table 1 [26],

where:

1. True Negative Average (TNA): Represents the number of wrong capital letter images that has been incorrectly categorized.
2. False Positive Average (FPA): Represents the number of wrong capital letter images that has been correctly categorized:

$$FPA = \frac{FPA}{(FPA + TNA)} * 100 \tag{6.8}$$

3. True Positive Average (TPA): Represents the number of right capital let-

ter images that has been correctly categorized.
4. False Negative Average (FNA): Represents the number of right capital letter images that has been incorrectly categorized:

$$FNA = \frac{FNA}{(FNA + TPA)} * 100 \tag{6.9}$$

5. Negative Predictive Average (NPA): which represents the negative prediction rate of the system.

$$NPA = \frac{TNA}{(TNA + FNA)} * 100 \tag{6.10}$$

6. Recall: Represents the scale sensitivity and the system's ability to find images where the wrong characters:
7. Precision(Pre): Represents the positive prediction average of the system.

$$Pre = \frac{TPA}{(TPA + FPA)} * 100 \tag{6.11}$$

8. Specificity(Spe) : It represents the system's ability to find the correct capital letters images.

$$Spe = \frac{TNA}{(TNA + FPA)} * 100 \tag{6.12}$$

From the results in Table 1, we note that the values of DR and Recall are very high, Where the ratio of DR is 97.5 while the ratio of Recall is 90 percent which means that the proposed algorithm was able to distinguish the capital letters in English efficiently and the performance of the proposed algorithm to retrieve the target goal and locate it among several different goals is very high. It is a successful algorithm to move from exploration to exploitation in a large way. We also note that the ratio of FR was 2.5, which is a relatively small percentage meaning that the system was unable to distinguish this low percentage from images of English letters which may be difficult to distinguish in all cases and there is no defect in the proposed system as they are distorted or omitted with very strange letters. The result of the Spe and Pre scales in the same table is a high result (90 percent) which means that the system has a high ability to sense images that match the type of image entered and it is also largely able to exclude images that contain types contrary to the type of image entered which indicates the high efficiency of the proposed system. While the high percentage NPA also that

Table 2: A comparison of Results With Other Algorithms

No.	**Algorithm used**	**Rate (percent)**
1	Bayesian decision theory [8]	92
2	Artificial Neural Network (ANN) with single output neuron [9]	92.59
3	Artificial Neural Network (ANN) with two hidden layers [10]	82.5
4	Multilayer perceptron neural network with a single hidden layer [11]	95
5	New Particle Swarm Optimization algorithm (NPSO)	97.5

appears in the table with a value of 90 percent which is a decent percentage for this system meaning that the proposed system has a high ability to predict negative goals which are excluded from among the promising goals and discover strange images about database images that do not return to the capital English letters. While we note that the FNA and FPA ratios are not a few values (10 percent) because the number of images that were examined as strange on the database used is not a large number, we have these ratios. Despite this, the algorithm achieves promising results and distinguishes strange images, whether these images contain letters belonging to the group of letters used in the system or they are not affiliated with the group. A comparison has been made between the results of the NPSO algorithm and the outcome of a group of studies in the same field as shown in Table 2.

7 Conclusion and Future Work

Primary treatment steps help to increase the system accuracy by achieving a high recognition rate, where DR and Recall were 97.5 and 90 regarding the extent of the system's recognizable. However, the use of histogram equalization technique to change the Image brightness and increase its Contrast had an important impact in increasing the recognition process precision. Also, the FR measure was small referring to the capability of the scheme to commonly detect fake images. The system is effective in analyzing capital letter images that appear in various sizes and fonts. The input of capital letter image values that is primarily processed. Also, we introduced the NPSO algorithm as a new way to build a strong and robust tradition of PSO image data set results. NPSO has been able to find the best goals that are close to each other, we used the correlation formula to find fitness for this function, and from the results in the tables it will be clear that the algorithm will find the best goals in high efficiency and short time. We can develop this work in the future to be used with small and handwritten letters which consider a useful and advanced step in the analysis of natural language. Furthermore, for future work, it is possible to develop a research project recognizing a word instead of a letter.

References

[1] Rajul Jain, Nitin Pise, Feature selection for effective text classification using semantic information, International Journal of Computer Applications, **113,** no. 10, (2015).

[2] Aurangzeb Khan, Baharum Baharudin, Lam Hong Lee, Khairullah Khan, A review of machine learning algorithms for text-documents classification, Journal of advances in information technology, **1,** no. 1, (2010), 4–20.

[3] Min Song, Handbook of research on text and web mining technologies, 2008.

[4] Sean Luke, Essentials of Metaheuristics.

[5] Ehsan Ali Al-Zubaidi, Maad M. Mijwil, Aysar Sh Alsaadi, Two-Dimensional Optical Character Recognition of Mouse Drawn in Turkish Capital Letters Using Multi-Layer Perceptron Classification, Journal of Southwest Jiaotong University, **54,** no. 4, (2019).

[6] Veronica Ong, Derwin Suhartono, Using k-nearest neighbor in optical character recognition, ComTech: Computer, Mathematics and Engineering Applications, **7**, no. 1, (2016), 53–65.

[7] Faiz Alotaibi, Muhamad Taufik Abdullah, Rusli Bin Hj Abdullah, Rahmita Wirza Binti Ok Rahmat, Ibrahim Abaker Targio Hashem, Arun Kumar Sangaiah, Optical character recognition for quranic image similarity matching, IEEE Access, **6**, (2017), 554–562.

[8] Mujtaba Husnain, Shahid Naweed, English letter classification using bayesian decision theory and feature extraction using principal component analysis, European Journal of Scientific Research, **34**, no. 2, (2009), 196–203.

[9] Tirtharaj Dash, Tanistha Natak, English Character Recognition using Artificial Neural Network, arXiv preprint arXiv:1306.4621, (2013).

[10] Yusuf Perwej, Ashish Chaturvedi, Neural networks for handwritten English alphabet recognition, arXiv preprint arXiv:1205.3966, (2012).

[11] Iliyasu Adamu, Kabir Ismail Umar, Hillary Watsila, Bello Musa Ibrahim, Multi-Layer Perceptron Network For English Character Recognition

[12] Rafael C. Gonzalez, Richard Eugene Woods, Steven L. Eddins, Digital image processing using MATLAB, 2004.

[13] Diya Chudasama, Tanvi Patel, Shubham Joshi, Ghanshyam I. Prajapati, Image segmentation using morphological operations, International Journal of Computer Applications, **117**, no. 18, (2015).

[14] Satish Pawar, V. K. Banga, Morphology Approach in Image Processing, International Conference on Intelligent Computational Systems, Dubai, (2012), 148–150.

[15] Vassilios Mardiris, Vassilios Chatzis, A Configurable Design for Morphological Erosion and Dilation Operations in Image Processing using Quantum-dot Cellular Automata, Journal of Engineering Science & Technology Review, **9**, no. 2, (2016).

[16] Nidhal Khdhair El Abbadi, Enas Hamood Al Saadi, Blood vessels extraction using mathematical morphology, Journal of Computer Science, **9**, no. 10, (2013), 13–89.

[17] Yadav Abhishak, Yadav Poonam, Digital image Processing, 2009.

[18] Omprakash Patel, Yogendra Maravi, Sanjeev Sharma, A comparative study of histogram equalization based image enhancement techniques for brightness preservation and contrast enhancement, arXiv preprint arXiv:1311.4033, (2013).

[19] Stan Birchfield, Image processing and analysis, 2016.

[20] Shanthi Muthuswamy, Sarah S. Lam, Discrete particle swarm optimization for the team orienteering problem, Memetic Computing, **3**, no. 4, (2011), 287–303.

[21] Yudong Zhang, Lenan Wu, A hybrid TS-PSO optimization algorithm, Journal of Convergence Information Technology, **6**, no. 5, (2011), 169–174,

[22] Basanagouda Pati, S. B. Karajgi, Optimized placement of multiple FACTS devices using PSO and CSA algorithms, International Journal of Electrical & Computer Engineering, **10**, no. 4, (2020), 3350–3357.

[23] Ekhlas Abbas Albahrani, A new audio encryption algorithm based on chaotic block cipher, Annual Conference on New Trends in Information & Communications Technology Applications, IEEE, (2017), 22–27.

[24] Ansam Nizar Younis, Fawzia Mahmoud Remo, Distinguish Musical Symbol Printed using the Linear Discriminant Analysis LDA and Similarity Scale, International Journal of Computer Applications, **975**, 8887.

[25] Ansam Nazar Younis, Developing and Implementation of Metaheuristic Algorithm Sequentially and Parallel for Musical Notes Recognition, University of Mosul, 2018,

[26] Ansam Nazar Younis, Fawziya Mahmood Ramo, A new parallel bat algorithm for musical note recognition, International Journal of Electrical and Computer Engineering, **11**, no. 1, (2021), 558–566.