

Evaluation of Theory based Handwritten Answers through BFO Model for Pixels and Pruned Scale Invariant Character Features

Rajesh Agrawal

*School of Computer and Systems Sciences, Jaipur National University, India and *Infosys Limited, Jaipur, India,*

Harvir Singh

School of Engineering and Technology, Jaipur National University, Jaipur, India

Vijay Singh Rathore

Dept. of Computer Science and Information Technology, IIS University Jaipur, India

Saurabh Maheshwari

Dept. of Computer Engineering, Govt. Women Engineering College Ajmer, India

correspondence: ¹rajesh.agrawal1977@gmail.com, ⁴saurabh@gwecca.ac.in

Abstract

Handwritten theory examinations are essential but they come with the complexity of evaluation by subject experts. It requires time consuming efforts from expert evaluators to check answer sheets of many students. Automation has been brought by evaluation of scanned sheets manually by evaluator through an online portal still the effort of evaluator is same and sometimes even more difficult due to problems associated with screen reading. This complete system needs automation such that the system has skills and intelligence equivalent to an expert evaluator to generate scores for the specific subjects. It is essential to recognize handwritings of thousands of candidates, each having unique features. The proposed method has thought of an innovative technique to train the system for every subject and also evaluate the paragraph answer written by the subjects. Segmentation of characters from continuous handwritten text has been done through a novel method inspired from Bacteria Foraging optimization (BFO). The BFO based pixel model spreads bacterial colonies over the text. Healthy colonies are used to identify valid characters while the unhealthy colonies are eliminated. The offspring bacterial colonies produce optimal characters. Final colony arrangements are compared with each other for character recognition through proposed pruned scale invariant features-based method. Handwritten text obtained from standard dataset for more than 50 subjects have been segmented, recognized and scored with optimistic accuracy comparative to the prevalent handwriting recognition methods.

Index Terms: *Handwritten text recognition, character segmentation, character recognition, SIFT, examination evaluation, scanned answer sheet evaluation.*

1. INTRODUCTION

Handwritten Examinations are essential to evaluate the analytical skills and to examine whether the candidate can hold on nerves in tight situations. On the other side evaluation of writing skills is equally important. But these lengthy theory examinations come with complexity of evaluation by subject experts. Lot many hours are devoted by academicians in answer scripts' evaluation and marking. It requires tedious effort from expert evaluators to check answer sheets of many students. This is time consuming process and it commonly happens that the expert evaluators may not take much interest in the work. In developed

countries, the online typing-based examinations are taking place of handwritten examinations but in developing and underdeveloped countries, the keyboard typing based test may not be feasible due to lack of computer knowledge and typing efficiency. Many people insist that they feel comfortable while writing handwritten answers instead of typing them. Considering the importance of handwritten answers automated evaluation is essential. Universities can initiate evaluation of the answer scripts by providing scanned copies of the scripts to the evaluators through an online portal. This may decrease the time of evaluation of answer script since the method can be executed online at remote locations maintaining anonymity and secrecy but the process of evaluation of scanned answer scripts is still lengthy and tiresome for evaluator as he has to evaluate each scanned copy in detail from the screen. Many people are still not comfortable in screen reading. Automation in this process is needed to eliminate need of examiner for automated evaluation and scoring of scanned documents. The solution is to recognize the text written in the answer script and convert it into digital script. But only conversion is not enough, the identification of the correctness of answers require intelligence of evaluator in the automation algorithm. The handwritten character (HC) and word recognition is in practice for a long time and methods such as [1] [2] [3] which have shown decent accuracy. Our objective is to enhance the applicability of these methods for continuous handwritten text such as theoretical answers written during examinations. Generally, a window of constant or adaptive size is used to scan the text by moving in the direction of writing. It is done to extract lines from the paragraphs. Content of each of the moving windows are analysed to extract specific features and matched with features of the stored character in the character library, this also includes pixel by pixel matching [4] [5] [6]. Constant size window faces a problem of improper coverage of a written character as it may leave some part of oversized characters while small sized characters may cause occurrence of some part of adjacent character in the candidate window. Authors in [7], surveys the idea of an adaptive window having dynamic window size but still window may have unwanted pixels. There are two problems, first, it is infeasible to cover a character optimally by a single window and second, moving windows spanning all the pixels horizontally have very high redundancy. We propose an innovative method to generate large variety of segments by bacteria foraging optimization model on pixels. We consider bacteria to be deployed at a black pixel where some text is written, serving as bacteria food or health function. The bacteria are arranged along the darker pixels of the text to form bacteria colonies of specific shapes and sizes thus localize characters in the text. Generation of adaptive character segments is done by evaluation of vertical bacterial saturation. The unhealthy colonies are eliminated. We use reproductions and dispersion steps of bacteria foraging to generate next generation segments which are combinations of bacteria colony shapes. The offspring bacterial colonies produce optimal characters. Final colony arrangements are compared with each other for character recognition through Scale invariant features. The work presented in this paper has three phases, first is to segment individual characters from continuous handwritten texts, second is to identify the characters and combine them to form words, i.e. to digitalize the answer sheets, and finally, we evaluate the answer given by the writer by matching the recognized words from the keyword based standard answers.

2. LITERATURE SURVEY

In this section, we discuss the status of automation in this field. Various steps of HWR. Detailing of segmentation and feature extraction methods has been done to identify the scope of improvements. Then we formulate the problem and suggest solutions for the problems.

2.1. Steps of HWR

HWR includes following steps which are database collection, image acquisition, skeleton detection, sub-word detection, feature extraction, training and testing of the system/recognition and post-processing. A database is collection of handwritten script and words. Some databases are IFN/ENT (Tunisian town and village names) [8], CVL (1 German and 6 English text) [9], IAM (collection of full English sentences) [10], ICDAR, RIMES [11] and PHOND. Databases have words and characters written by different writers and

is used for training, testing and validation of classifier model. We explain state-of-the-art segmentation and feature extraction methods here.

2.2. Segmentation Methods

Segmentation is performed at different levels. First is extracting word from the skeletonized text and then isolating characters from a word.

Junction Based Segmentation: Junction are the points where more than one stroke of the character meets. The characters may be uniquely determined on the basis of number of strokes detected. Rule-based segmentation algorithm decided for the specific junction locations is done as phase two of this segmentation algorithm completing the segmentation termed as full segmentation [12].

Baseline Pixel Burst Method (BPBM): The case where the characters are touching each other and joined is most difficult to segment specially for English language handwritten script. Authors [13] have proposed BPBM specifically for segmentation of the isolated characters cases.

Zone Wise Segmentation: Similar to vertical segmentation cuts between the characters there are different horizontal zones divided for the text towards Indic handwriting recognition [14]. The different areas are named as lower, middle and upper zone between the boundaries of baseline and Matra/Shirorekha space. The method is capable to divide the characters into unique disjoint classes of characters which can be individually now classified for unique identification [14].

Region Based Segmentation: Till now touching characters segmentation is not solved with higher accuracy. An innovative method for offline cursive words segmentation having touching characters has been tried to solve promisingly in [15] based on region-wise segmentation.

Component Based Segmentation: Fundamental principle of such methods is to segment various components out of the running text. Histogram projection analysis and regrouping of the components is utilized [16].

2.3. Feature Extraction Methods

From the extracted segments, various features are extracted which serves as unique criteria to recognize the text. The features portray either the shape, size of character or the way of writing of each author. Some of the main feature extraction methods can be classified as below:

Character Structural Features: The structural information of the script such as stroke size, no. of bifurcation points, contours and no of circles serves as unique structural features of the script.

Topological characteristics: This class of features are used to build a character using different types of strokes, cross points, loops, end of line and branch point.

Geometrical features: The curvature of character, word length, along with the fraction of height and width of character bounding box, distance between ending point and y-min vertically, distance between initial and ending point vertically or horizontally, distance covered by two points, words lower and upper mass, stroke width and strokes distance are some of main features described [17].

Statistical Feature: The main utilization of these features is to understand the writing style variations among the authors. Statistical methods are used to reduce the dimension of the feature sets [18].

Pyramid Histogram of Oriented Gradient (PHOG) Feature: A window slides over the text to estimate the spatial shape features of the script via layout and shape of the component inside the window. A gradient orientation is decided at each level of pyramid which is used to separate contents of the sliding windows and convert them into cells [14].

2.4. Compelling Problem: Paragraph Based Answer Evaluation

Till now we have seen the common handwriting recognition problems, applications and generic method. Lengthy theory examinations come with complexity of evaluation by subject experts. Lot many hours are devoted by academicians in answer scripts' evaluation and marking. It requires tedious effort from expert evaluators to check answer sheets of many students. This is time consuming process and it commonly happens that the expert evaluators may not take much interest in the work.

In developed countries, the online typing-based examinations are taking place of handwritten examinations but in developing and underdeveloped countries, the keyboard typing based test may not be feasible due to lack of computer knowledge and typing efficiency. Many people insist that they feel comfortable while writing handwritten answers instead of typing them. Considering the importance of handwritten answers automated evaluation is essential. Universities can initiate evaluation of the answer scripts by providing scanned copies of the scripts to the evaluators through an online portal. This may decrease the time of evaluation of answer script since the method can be executed online at remote locations maintaining anonymity and secrecy but the process of evaluation of scanned answer scripts is still lengthy and tiresome for evaluator as he has to evaluate each scanned copy in detail from the screen. Many people are still not comfortable in screen reading. Automation in this process is needed to eliminate need of examiner for automated evaluation and scoring of scanned documents.

Following are the challenges to be addressed in this research:

1. It has been observed that the recognition of cursive or running handwritings is more difficult since it involves overlapping and uneven distances among the characters. This reduces the identification accuracy.
2. The characters are usually are of different sizes which cannot fit in a constant size window [19]. So, window based cutting methods are not effective.
3. Approaches have worked to segment character by cutting where the pixel density is minimum, considering it to be a connector between the words. This can lead to false cuttings for characters like w, m, n, h, u, v & y, where a false cut can take place inside the characters considering the shape of character as a connection [20].
4. Combinations of the characters such as r and n can lead to confusion as m
5. Using a standard shape character library and their shape features may not work for all subjects which have deformed handwriting and do not make symbols in standard ways. So personal training is essential single model trained for some subjects cannot be extended for other subjects.
6. A common method is needed to work on different language scripts which may have different character sets and different writing direction.
7. After identification of the words, the task of evaluation of the answers is there which require intelligence of the examiner.

After recognition we use keyword-based examination evaluation which has been used for sentimental analysis field for identification of the sentiments based on availability of predefined keywords. We use it to evaluate the score of examination based on keywords found in the recognized text.

3. PROPOSED WORK

This paper has an objective to develop a method for digitalization of handwritten text-based paragraph answers. The proposed method for the answer sheets evaluation is divided into three phases First training for handwriting features of each candidate appearing in the examination, second, digitalization of the handwritten words in the scanned answer sheets of the candidate and third phase is to evaluate the scanned answer sheet of the candidate and estimate marks on the basis of keywords.

We make following assumptions to ensure the accuracy for our proposed system. First, the model requires to be trained for each subject to recognize unique handwritings. Second, the Subject is asked to write a

paragraph which has been custom designed to encompass the characters set of the target language. Third, results from previous methods can be compared by implementing the proposed algorithm over standard datasets. Currently, only 1 question has been answered by the subject, so one evaluation per subject. This can easily be extended by incorporating each new answer on a new page which may have an optically verifiable code. Fourth, the subjects are not permitted to draw characters in completely different manner intentionally during examination. Fifth, each page has a unique identity and start of answering a question is to be done from question specific page, currently we take the image file name as the subject and question ID.

3.1. Development of the solution

The input is fed in the form of scanned image of handwritten texts obtained in the form of answer to a question in an examination. These scanned images are pre-processed to enhance the target information through contrast enhanced and binarization. This may be followed by thinning depending upon stroke width, some famous standards for these procedures have been adopted from [21]. Apart from this, different handwritings have different angles of inclination which has been normalized by using method proposed in [22].

3.2. Mapping of Bacteria Foraging Optimization on Character Image Segmentation

This section has been dedicated to map character segmentation to the bacteria foraging method of swarm intelligence. It is a bio-mimic algorithm meant for global optimization of a system. The method works on mathematical modelling of foraging of Bacteria. Case study of bacteria E.coli is popular which performs local actions to obtain a global solution of finding for optimal food quantity. The genetic variation obtained through reproduction help either to better offspring or discarding of bacteria reaching to non-optimal food locations. The nature plays role as optimizer by selection and through reproduction for future generations in such a way that over-a-time-period the target bacterial colonies which are healthy above desired threshold only move to offspring generation and unhealthy or undesirable bacteria not having enough food will get eliminated.

3.3. Chemo-taxis and Tumbling

A line and word can be easily extracted from the paragraph text. Being continuous in nature, it is difficult to extract characters as discussed earlier, through moving windows. We envision to use the Bacterial foraging algorithm (BFO) [23] for the target segmentation. The idea is to assume that the bacteria will grow in the image of the target text and the black pixels denoting the text is food of the bacteria over which bacteria spread. It has been assumed that the food function of bacteria is dependent upon the intensity of the pixel, with maximum food available when pixel intensity is zero. Following is the optimization food function.

$$F(x,y) = \begin{cases} \text{if } I_m(x,y) == 0 \text{ then } F = 1 \text{ perform chemtaxis} \\ \text{if } I_m(x,y) == 1 \text{ then } F = 0 \text{ perform tumbling} \end{cases}$$

The deployment of bacteria is based on the objective food function while following chemotaxis step. The value of F is 1 when pixel intensity is zero thus denoting the maximum food location for the bacteria. The bacteria have natural tendency to repel from the pixel locations where value of F is less than 1 i.e. locations where pixel intensity is not black during the tumbling step. This spread of the bacteria takes over entire image and start consuming food. The bacteria collections are divided into first level related colonies. Each of these denotes a line of text. The low density of bacteria draws boundary of colonies and serve as the basis of segmentation. A high saturation of bacteria in a linear form across horizontal axis denotes text line and vice versa is used as line boundaries.

Similarly, low saturation along vertical axis of the image draws second level of related colonies denoted

candidate words. The vertical saturation of bacteria decreases at the boundaries of the words which can be used to cut down the text into words by choosing a minimum threshold of bacterial minima. The same method can be extended to cut down characters from the segmented words but clear boundaries are not available to segment out the characters from the words specially in the case of continuous handwritings where there is no break between the characters. Vertical separations of bacterial colonies are difficult to estimate due to existence of bacteria all over so we propose a dynamic bacteria colony cutting method which scans each word and analyze the density of bacterial colonies from left to right or in the direction of writing. During these scans some specific conditions are observed which are as follows:

1. The connection between two adjacent characters are indicated by bacteria density lower than a threshold such as between 'a', 't' and 'g' in figure 1.
2. There can be complete disconnection between any two adjacent characters distinguished with help of density lower than a threshold as shown in figure 1 for characters 'm' & 'i' and 'i' & 'n'.
3. There can be density lower than a threshold due to variation in stroke also such as cut inside the character 'o' and 'a'.
4. The most important observation is for those characters which have twice descents such as 'w', 'm', 'n', 'h', 'u', 'v' and 'y'. the cutting on the basis of threshold leads to a false cut inside the characters. This is due to algorithm behaving to confuse the descent inside the character as connection between two characters.

The cuts consisting of true and false are shown in figure 1 where type 1 and 2 cuts are target true cuts but type 3 and 4 are false cuts which will exist while separation. All the four types of cuts are our candidate cuts for character segmentation or so called first generation bacterial colonies. It is essential to overcome false cuts.

3.4. Reproduction, Dispersion and Elimination

Candidate segments generated from the character cuts are fed to further steps of BFO algorithm to remove false cuts and interpret correct character segments. All the candidate bacterial colonies of first generation have been generated by their property of showing probable disconnection as shown in figure 2. The false cuts may have led to improper disconnection of type 3 or 4. It means that combination of two nearby cuts may produce a true character segment. This is done through reproduction step of BFO where three adjacent colonies combine to make possible offspring.

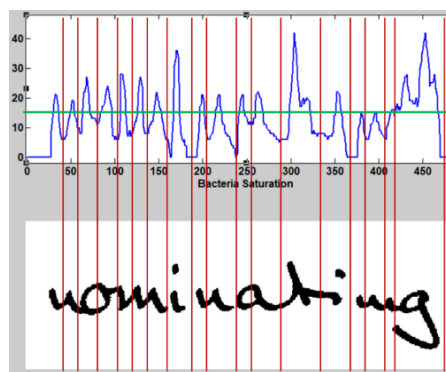


Fig 1: Cutting Character Segments by thresholding the vertical



Fig 2: Segmented bacterial colonies

Fig 3: Process of reproduction

First offspring is obtained by reproduction of any colony C_i and C_{i+1} and then another is obtained by C_i , C_{i+1} and C_{i+2} . These colonies are further candidate colonies for true character segments. In this way all the false cuts not interpret the correct character at least with any one or more offspring. The process of

reproduction is shown in figure 3. Due to reproduction new offspring colonies generated are shown in figure 4, it may be noted that these have all the characters form the words in segmented form. Further, the candidate bacterial segments are examined for correctness by matching them with stored segments during the training phase through methods such as block based, SIFT feature or visual correction. During dispersion step of foraging, this matching take place and filtration of true segments is performed and their labels are recorded. These chosen colonies are healthy enough to generate next generation recombination i.e. a word. This phase also causes elimination action on the candidate colonies which does not match with any of the stored characters and are not fit enough to survive for next generation. All the offspring colonies will now produce words.



Fig. 4: offspring colonies after reproduction step

3.5. Production of Characters from Bacterial Colonies

Healthy and optimal offspring are generated with the help of BFO while others have been eliminated. All the survived character segments now have a character label associated with them based on their shape and size features. The next generation word is generated by combining all of these character labels in the direction of writing. The same process can be used for any of the directions starting from top or left or vice versa.

3.6. Feature Extraction

In this section we explain the features extracted from the candidate bacterial colonies deployment shapes. We observed that the candidate bacterial arrangements have specific shapes. We use following features.

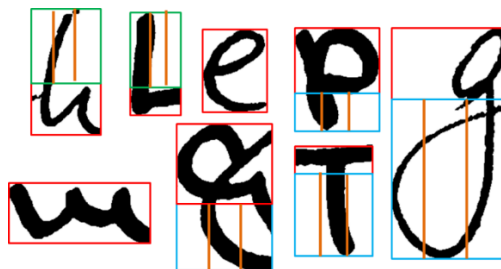


Fig. 5: Horizontal Segments- Lower and upper low saturation segments are divided into overlapping horizontal segments shown encircled by brown boxes.

3.6.1. Overlapping zones-based features

Divide the character into higher saturation area and above and below lower saturation areas. The higher saturation area is candidate for major shape features so used completely while lower saturation areas are divided into sub segments. E.g. English language lower case characters such as 'b', 'd', 'h', 'k', 'l', 'p' and 't' have specific lines or curves either in upper right half or left half while characters 'g', 'j', 'p', 'q', 'y', 'z', have specific shapes in lower part of the character. 'f' is a lower-case character which may have specific shapes in both upper and lower section. Remaining characters have major shape features in the center part

of the bacteria colony. The uppercase characters have unique distribution throughout the character, so they can be better identified through distribution-based features. In this way, vertical zones in the character are created while considering horizontal saturation of bacteria. Character specific features can also be extracted from horizontal zones based vertical bacteria saturation. The horizontal and vertical zones are shown in figure 5. For the trade-off between the number of vertical zones and computation complexity we choose 3 vertical zones per character to be created, these will overlap to the horizontal zones. These overlapping zones we used to cut the character to estimate feature. One of the features is the bacteria saturation feature which is estimated as follows:

$$BS = \frac{\sum \text{No. of Bacteria Present in a colony}}{\text{Area of Colony}}$$

The area of colony is the area of the respective segment. Considering the size of any image segment to be S rows and T columns so the area of the segment is SxT.

Figure 5 shows high saturation segment encircled by red box and upper low saturation segment by green box and lower low saturation segment is encircled by blue box. This is how each character is divided into vertical zones using horizontal saturation.

3.6.2. Local Descriptors

Local descriptor such as Scale Invariant feature transform (SIFT) descriptors can estimate the scaling and rotation invariant features in the presence of noise. We believe that the variations in the handwriting such as change of orientation, bigger/smaller characters than usual and some handwriting effects due to fast writing can be eliminated using this feature. We estimate SIFT descriptors [24] of complete characters. Following figure 6 shows different scaled and rotated 'e', 'M' and 'w' characters from the same writers. The SIFT features have been obtained for all the three. It can be seen that the SIFT matching is with the corresponding character points which may be anywhere in the image. Slight shape and size changes as in the characters 'e' and 'M' by the same writer can be easily identified with SIFT descriptor features while rotated character 'w' written by same writer also have matching SIFT descriptors. The blue line segments are connecting the matching descriptors. During Matching of the features, the corresponding segment features are compared with each other. The matching features are reported by SIFT matcher algorithm [24] [017], on the basis of the number of features matching and corresponding score, we estimate a dynamic descriptor matching score. So, it is necessary to estimate all kinds of features.

3.6.3. Descriptors Matching Score

Geometry of bacteria colony after optimization is treated as a pattern and scale invariant features are used to compare scale invariant features of bacteria colony. Let us suppose we need to match two segments S1 and S2 of different sizes. First, we pad the smaller segment and make the length the breadth of the image same. The segments are now fed to SIFT feature extractor algorithm [24]. This return SIFT feature descriptors. These descriptors are then input to matching algorithm which returns the matching pairs of descriptors and corresponding scores calculated by the squared Euclidean distance between the matches. To ensure that any two matching pairs are not spatially away from each other, we put a threshold over the score. The matches with the score below threshold are only considered for estimation of distance. Mean of all distance scores is used as final matching score.



Fig. 6: (A) SIFT key points and matching, blue lines showing key points correspondence (B) Pruned SIFT key points and matching, blue lines showing key points correspondence

3.6.4. SIFT Matching points pruning algorithm

We noticed that all the SIFT matching points do not belong to the visually correct matches for some areas of the character as shown in figure 6. We investigated and found out that sometimes more than one feature descriptor of first image matches with multiple descriptors of second image. This causes a confusing matching score. We are hereby proposing a matching descriptor pruning algorithm. We collect all the descriptors of first image which match with multiple descriptor of second image. Among them we discard all matching pairs except the one with maximum scores. This reduces the effective number of matches but results into true matches. Further, resultant matching score is estimated using pruned matched points and score. The pruned SIFT matching points for matching of figure 6 are shown in figure 6(B), it is clearly visible that the false descriptor matching are eliminated.

$$Dist_{olp,i} = (BS_{X,i} - BS_{Y,i}) \forall i = 1 \text{ to } 8$$

$$SC_{olp} = \sum_{i=1}^8 \begin{cases} \text{if } Dist_{olp,i} \leq th_i \text{ then output } 1 \\ \text{if } Dist_{olp,i} > th_i \text{ then output } 0 \end{cases}$$

$$SC_{SIFT} = \text{count}(\text{Match}(STFT(X), SIFT(Y)))$$

$$SC_{res} = \alpha \cdot SC_{olp} + \beta \cdot SC_{SIFT}$$

3.6.5. Resultant Score

This resultant score is estimated to combine the overlapping zones features and local descriptors features. The difference between two-character segments X and Y is estimated using the difference between the overlapping zones bacteria saturation and amount of matching. The score based on overlapping blocks are estimated using equation where Dist_i estimates the absolute bacteria saturation difference between the segments estimated as per equation and th_i corresponds to threshold per block for i=1 to 8. The summation of the output decision in equation leads to final score for overlapping blocks features. Similarly, the score is calculated for SIFT matching by counting the number of pruned matches when SIFT matcher algorithm is implemented between the two segments. SC_{res} is the resultant score from the two sub-scores with individual gain as α and β whose values have been determined empirically through various experiments. Currently, the results shown are for values 100 and 1 respectively. The match is declared for the database vector which matches with maximum score above threshold.

3.7. Keyword based Answer Evaluation

The first phase of the proposed system is to segment the characters out of a word, second phase is to recognize the segments and collect them as words. The final step is to evaluate and examination by the marking on the basis of detected keywords from the answer. The keyword dictionary is used for examination evaluation. This dictionary is list of keywords per answer which are necessary and sufficient to evaluate and answer. Each keyword has some weightage marks. The number of keywords recognized from the text is converted to the score. We manually generated 10 tests each written by different subject. Each test has list of 10 keywords to be found in the text written by the writer. Each keyword i from 1 to k of test t has weight K_{ti} . Weights of all the recognized words are summed up to get evaluation score. We also have the ground truth regarding the keywords available in the text. So, accuracy of examination evaluation can be obtained from here. The evaluation score is estimated using following equation.

$$EVS_t = \sum_{i=1}^k K_{ti} \text{ for test } t \text{ and keyword } i$$

The difference between the estimated evaluation score and ground truth for the query test is estimated, this is used as Evaluation Error estimate.

$$EEE_{ts} = \text{Actual Score}_{ts} - \text{Estimated Score}_{ts}$$

for test t taken by subject s

4. IMPLEMENTATION RESULTS

4.1. Implementation Tool and Dataset

The proposed algorithm has been implemented on MATLAB v14, image processing toolbox. The tasks of binarization, thinning, etc. can be done easily by using inbuilt methods while it is easy to implement logics for proposed segmentation and feature matching through custom programming support of MATLAB. Standard SIFT implementation MATLAB script by [24] has been used for SIFT feature extraction and matching. We designed the SIFT points running algorithm. MATLAB has another advantage of visualization of the results, plots, figures in effective manner. The simulation also has capability to check the value of any runtime variable, debug the script and many other programming related features making MATLAB an obvious choice for implementation of logic in the field of image processing and many other applications. We conduct experiments on the IAM Handwriting Database version 3.0 [6]. The dataset comprises of different writers writing paragraph text. The text is written on white paper and scanned at a resolution of 300 dpi. The dataset has advantage that the ground truth words for all the handwritten words in the texts have been given separately. This makes easier to verify the recognized word by matching it with corresponding ground truth. The dataset also has segmented words, which we directly use to apply proposed method and the name of the word is used as to find the ground truth word from the provided text files with respect to handwritten word. We use texts written by 50 writers for training and testing. Total number of words used in training is approximately 50 per subject having English alphabet. The system is trained with approximately 2500 words. We test more than 1000 words written by the same writers. We match only the features of the characters of the specific to author to match with query word. This is essential for the objective of examination evaluation. We concentrate on recognition so we directly used word segmented images. The ground truth available for all the words is used for character and word recognition accuracy as well as for false recognitions.

4.2. Processing Steps

Segmentation: The segmented characters are stored in the form of feature vectors for all the overlapping blocks. The statistical features, SIFT descriptors are stored in the database during the training phase.

Recognition: During query, input text is also segmented through proposed bacteria foraging algorithm. Each of the segments is compared with the stored feature vectors of all the characters of the same subject. Matching score is estimated. The stored character label is used as recognized character which has the maximum matching score. The estimated word is further used for keyword-based evaluation method. Following table 1 shows the recognized words versus ground truth words of a random sample of the words tested for a writer. It can be seen that all the words have been correctly recognized except in word ‘majority’ where ‘y’ and in word ‘resolution’ where ‘l’ have not been recognized.

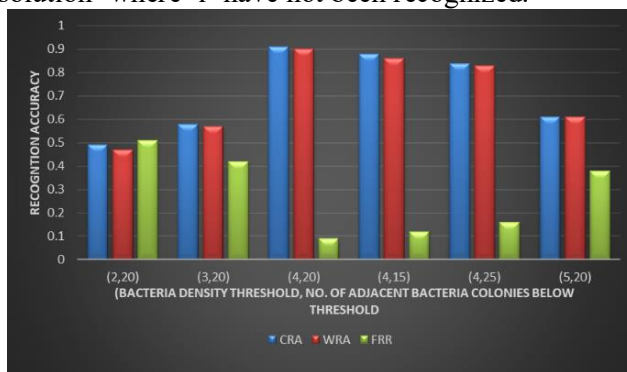


Fig 7: Values of recognition accuracies CRA, WRA and FRR for selection of optimal values of Bacteria Density Threshold and number of adjacent Bacteria Colonies below threshold

4.3. Bacterial Density Threshold Selection

The performance results shown earlier have been obtained on specific values of Bacteria Density Threshold and number of adjacent Bacteria Colonies below threshold. The bacterial density threshold decides the probability of completion of a character and measured in terms of number of bacteria maximum available in a column. But single colony having population less than threshold may not be chosen for character separator since it may be due to writing errors, so we choose a number of adjacent bacterial colonies below the density threshold to select the character cuts. The selection of bacterial density threshold has been done on the basis of the results of the performance parameters. The graphical comparison shown in figure 7 and 8 makes it easier to select the threshold for bacterial density as 4 and no. of adjacent colonies below threshold to be 20. Performance for other values of these parameters is less compared on the selected values.

4.4. Examination Evaluation

We estimate the evaluation score for 10 tests written by different writers by matching the recognized words from the keyword set of the answer. Corresponding score of the keywords found in the recognized text is summed up to get the score achieved by the subject in the test. We already estimate the actual score for the test manually. The difference between the estimated and actual score gives the evaluation error rate. The average EER is 2.2 marks out of 30 marks tests and the number of keywords per test is 10. Different keywords have different score weights.

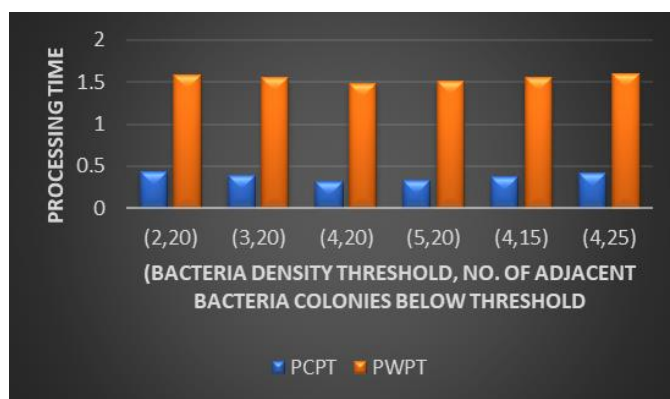


Fig 8: Values of processing times PCPT and PWPT for selection of optimal values of Bacteria Density Threshold and number of adjacent Bacteria Colonies below threshold

4.5. Analysis Parameters

1. **Character recognition accuracy**

Amount of characters recognized correctly individually from the available number of characters to be recognized

$$CRR = \frac{\sum_1^W \text{No. of Characters Recognized Correctly}}{\sum_1^W \text{No. of Characters in a Word}} \quad \forall \text{ words } 1 \text{ to } W$$

2. **Word recognition accuracy**

This is estimated as summation of the number of characters recognized correctly per word. The CRR and WRR have been reported in table 2.

$$WRR = \sum_1^W \frac{\text{No. of Characters Recognized Correctly}}{\text{No. of Characters in a Word}} \quad \forall \text{ words } 1 \text{ to } W$$

3. **Per Character Processing Time** It is estimated as the time required in extracting features of the character segment and matching with the stored features. The features include overlapping blocks features and the SIFT descriptors
- $$PCPT = \frac{\sum \text{time taken per character}}{\text{total number of characters}}$$
- The average processing time needed per character for extraction of feature and recognition is 0.3 sec.
4. **Per Word Processing Time** Time required to segment the word into characters, estimate features, and recognize the characters and combine them to recognize the words from ground truth.
- $$PWPT = \frac{\sum \text{time taken per word}}{\text{total number of words}}$$
- Per word processing time is needed for segmentation and recognition of a word is appx. 1.5 sec.

4.6. Computational Complexity

The time complexity has been identified in terms of real time execution of the proposed algorithm for recognition from PCPT and PWPT. The computational complexity can be measured as the size of the image taken input and number of times the image is scanned by the algorithm for segmentation and recognition tasks. Since our algorithm scans the image only once in comparison to previous approaches where the image is scanned multiple times for segmentation. In HOG based methods [19] the image is scanned pixel by pixel and variable or fixed size window is used to cut the segments, thus, the total number of segments created is approximately equal to the width of the image, this causes high proportions of overlapping and redundancy in the segments formed whereas in our case no two consecutive segments share any overlapping area so the number of final segments are less than the HOG based methods. Although the overlapping is essential for some characters identifications such as ‘m’, ‘n’, etc. which are compensated by combination of bacteria colonies during the reproduction step. The feature extraction of a segmented character has the complexity of creating overlapping blocks and SIFT [24] which is not significant for smaller size segments and binary images. It will also comprise the complexities of individual feature extraction algorithms used to make the final character template.

Recognized Word	Ground Truth Words
the	The
subject	Subject
and	and
he	he
be	be
by	by
Griffitus	Griffitus
majorit	majority
Labour	Labour
are	are
down	down
resoution	resolution

Table 1. Recognized versus ground truth words

Method	WRA %	CRA %
Static Window based Cut and HMM based recognition [9]*	86	90
Dynamic Size Cuts and one to one matching	50	51
Sliding window and HMM features [10]	55	NR
Sliding window and HMM/ANN features [25]	85.50	93.10
SIFT Features with spatial pyramids [26]	80	89.73
Proposed Dynamic Size cuts with Bacterial Colony optimized and pruned SIFT features	91.78	92.97

*on different dataset, NR-Not reported

Table 2. Comparison of proposed algorithm implementation results with recent approaches

4.7. Comparison with Previous Methods

4.7.1. Static Window based cuts and Human Markov Models (HMM)

In this class of methods, a static or dynamic window of one or more columns width moves along writing direction for each line. The height of window is equal to height of line. The HMM based statistical language model proposed has average recognition accuracy of 62% [10]. An HMM-based recognition system for Latin script to convert it into Arabic script has been proposed in [9]. HMM model features are used to uniquely identify character. The average best recognition rate of this system is 86%.

4.7.2. SIFT based methods

Use of SIFT can be seen for writer identification in [27, 28]. Combination of K-Adjacent Segments, SURF, and Contour Gradient Descriptors has been used for the writer identification problem [28]. We recognize the characters and words individually while in [27, 28] the features have been extracted to uniquely identify the individual. The subject distinguish rate in both is more than 98%. SIFT features have been used for word spotting in [26] with average word and character recognition error rate to be 20% and 11% respectively. Application of SIFT descriptors in [26] are extracted for complete word whereas we extract it for each segmented character giving more localized description capability to SIFT. Also, this is for the first time SIFT features are being used after proposed SIFT key points pruning algorithm for handwritten text recognition.

5. CONCLUSION

In this paper, we have made propositions for recognition of handwritten paragraph text-based answers and estimated equivalent marks for the answers. The work has been accomplished in three phases. The first phase is segmentation where the characters are segmented from the continuous text through proposed bacteria foraging optimization algorithm, second phase is to match the query segmented characters from the stored character patterns of the same subject through proposed pruned SIFT feature matching technique. The last phase is to use the recognized words to evaluate the answers based on proposed keyword-based marking method and scores for the text is generated. The recognition accuracies show a promising aspect of the research while we also compare the results with prevalent methods. The processing time is estimated for the analysis of computation complexity. All the results are impressive and optimistic. This is for the first-time handwritten theory-based test has been evaluated on the basis of keyword-based scoring. The score is calculated for the recognized text by comparing with pre-set score per word. The evaluation error rate gives an estimate for the evaluation accuracy. We are able to generate score for various handwritten tests given by users with low EEE.

REFERENCES

1. A. Priya, S. Mishra, S. Raj, S. Mandal, and S. Datta, 2016, April. Online and offline character recognition: A survey. In Communication and Signal Processing (ICCSP), 2016 International Conference on (pp. 0967-0970). IEEE.
2. A. Lawgali, 2015. A survey on arabic character recognition. International Journal of Signal Processing, Image Processing and Pattern Recognition, 8(2), pp.401-426.
3. Q. Ye, and D. Doermann, 2015. Text detection and recognition in imagery: A survey. IEEE transactions on pattern analysis and machine intelligence, 37(7), pp.1480-1500.
4. N. Islam, Z. Islam, and N. Noor, 2017. A Survey on Optical Character Recognition System. arXiv preprint arXiv:1710.05703.
5. U. V. Marti and H. Bunke, 2000. Handwritten sentence recognition. In Pattern Recognition, 2000. Proceedings. 15th International Conference on (Vol. 3, pp. 463-466). IEEE.
6. U. V. Marti and H. Bunke, 2001. Text line segmentation and word recognition in a system for general writer independent handwriting recognition. In Document Analysis and Recognition, 2001. Proceedings. Sixth International Conference on (pp. 159-163). IEEE.
7. M. Rabi, M. Amrouch, & Z. Mahani (2017). A survey of contextual handwritten recognition systems based HMMs for cursive arabic and latin script. Int. J. Comput. Appl, 160(2), 31-37.
8. M. Pechwitz S. Snoussi V. Märgner, N. Ellouze and H. Amiri, "IFN/ENIT Database of Handwritten Arabic Words," in Proceeding of Francophone International Conference on writing and Document, Hammamet, pp. 127-136, 2002.
9. N. Arica and T. Fatos, Y. Vural, "An Overview of Character Recognition Focused on Off-Line Handwriting", IEEE Transactions on Systems, Man, and Cybernetics: Applications And Reviews, VOL. 31, NO. 2, MAY 2001.
10. U. V. Marti and H. Bunke, "The IAM-database: an English sentence database for offline handwriting recognition", IJDAR (2002), pp: 39-46.
11. F. Kleber, S. Fiel, M. Diem and R. Sablatnig, "CVL-DataBase: An Off-Line Database for Writer Retrieval, Writer Identification and Word Spotting," 2013 12th International Conference on Document Analysis and Recognition, Washington, DC, 2013, pp. 560-564.
12. Y. LU, "Machine Printed Character Segmentation--An Overview", Pattern Recognition. Vol. 28, No. 1, 1995, pp. 67-80.
13. A. Hamid and R. Haraty, "A Neuro-Heuristic Approach for Segmenting Handwritten Arabic Text", IEEE International Conference on Computer systems and Applications (AICCSA) 2001.
14. A. Choudharya, R. Rishib and S. Ahlawate, "Off-Line Handwritten Character Recognition using Features Extracted from Binarization Technique", AASRI Conference on Intelligent Systems and Control, 2013, pp. 306-312.
15. U. Dwivedi, P. Rajput, M. K. Sharma, & G. Noida (2017). Cursive Handwriting Recognition System Using Feature Extraction and Artificial Neural Network. International Research Journal of Engineering and Technology (IRJET), 4(03).
16. R. K. Mandal and N R Manna, "Handwritten English Character Segmentation by Baseline Pixel Burst Method (BPBM)", AMSE Journals –2014-Series: Advances B; Vol. 57; N° 1; pp 31-46.
17. F. Kurniawan, M. Shafry, D. Daman, A. Rehman, D. Mohamad and S. M. Shamsudin, "Region-Based Touched Character Segmentation in Handwritten Words", International Journal of Innovative Computing, Information and Control Volume 7, Number 6, June 2011.
18. H. A. Jawad, J. Jianmin, R. Jinchang, S. I. Stan, "Component-based Segmentation of Words from Handwritten Arabic Text", Proceedings of World Academy of Science, Engineering and Technology, Volume 31, pp 1307-6884, 2008.
19. R. Girshick, J. Donahue, T. Darrell, and J. Malik, 2016. Region-based convolutional networks for accurate object detection and segmentation. IEEE transactions on pattern analysis and machine intelligence, 38(1), pp.142-158.

20. P. M. Manwatkar, & K. R. Singh, (2015, January). A technical review on text recognition from images. In 2015 IEEE 9th International Conference on Intelligent Systems and Control (ISCO) (pp. 1-5). IEEE.
21. S. Fiel, and R. Sablatnig, 2012, March. Writer retrieval and writer identification using local features. In Document Analysis Systems (DAS), 2012 10th IAPR International Workshop on (pp. 145-149). IEEE.
22. N. Kato, M. Suzuki, S. I. Omachi, H. Aso, and Y. Nemoto, 1999. A handwritten character recognition system using directional element feature and asymmetric Mahalanobis distance. IEEE transactions on pattern analysis and machine intelligence, 21(3), pp.258-262.
23. Z. Zeng, L. Guan, Y. Zhu, Q. Liu, & J. He, (2017). A novel image segmentation method based on an improved bacterial foraging optimization algorithm. Journal of Information Hiding & Multimedia Signal Processing, 8(2), 348-357.
24. D. G. Lowe, "Distinctive image features from scale-invariant keypoints," International journal of computer vision, vol. 60, no. 2, pp. 91-110, 2004.
25. S. Espana-Boquera, M. J. Castro-Bleda, J. Gorbe-Moya, and F. Zamora-Martinez, 2011. Improving offline handwritten text recognition with hybrid HMM/ANN models. IEEE transactions on pattern analysis and machine intelligence, 33(4), pp.767-779.
26. J. Almazán, A. Gordo, A. Fornés, and E. Valveny, 2014. Word spotting and recognition with embedded attributes. IEEE transactions on pattern analysis and machine intelligence, 36(12), pp.2552-2566.
27. S. Fiel, and R. Sablatnig, 2012, March. Writer retrieval and writer identification using local features. In Document Analysis Systems (DAS), 2012 10th IAPR International Workshop on (pp. 145-149). IEEE.
28. G. Louloudis, N. Stamatopoulos, & B. Gatos, (2018). Writer Identification. Document Analysis And Text Recognition: Benchmarking State-of-the-art Systems, 82, 121.

Authors



Rajesh Agarwal is project manager at Infosys limited , Jaipur and currently working on assignmet at USA. He is having 17 years of experience in software development and maintenance with expertise in Project Management, Stakeholder Management, Design, Development, Maintenance and Implementation of systems in Microsoft technologies. Currently pursuing PhD from School of Computer and Systems Sciences, Jaipur National University, Jaipur, Rajasthan, India.
ORCID iD <https://orcid.org/0000-0002-9894-3490>



Harvir Singh Director & Professor at School of Engineering & Technology, Jaipur National University, Jagatpura, Jaipur. He has Teaching Experience of 24 Years. He has a member of various committees like Anti Ragging, Inspection, Sports council, examination and academic Council in Deenbandhu Chhotu Ram University of Science & Technology, Murthal (Sonapat).



Vijay Singh Rathore is currently Professor at Dept. Of Computer Science And IT, The IIS (Deemed to be) University, Jaipur. He has Teaching Experience of 18 Years. He is Past Chairman, CSI Jaipur Chapter, got Two Patents published, PhD Supervised (Awarded: 16, Under Supervision: 07), 77 Research Papers published (Includes Springer, Elsevier, IEEE, etc indexed with Scopus, Google, Thompson etc.). He has written 10 Books on various subjects of CSE, IT and Management and an active Reviewer in some National and International Journals.



Saurabh Maheshwari is currently working as Assistant Professor in Govt. Women Engineering College Ajmer, Rajasthan, India. He is currently pursuing PhD from IIT Jodhpur. He has done MTech from MNIT Jaipur.
ORCID iD <https://orcid.org/0000-0003-3638-7545>
ResearcherID: L-3115-2018
Scopus Author ID: 46661798600