



## A New Hybrid Algorithm for Telugu Word Retrieval and Recognition

Kesana Mohana Lakshmi<sup>1\*</sup>

Tummala Ranga Babu<sup>2</sup>

<sup>1</sup>*Chamakura Malla Reddy Technical Campus, Hyderabad, Telangana, India*

<sup>2</sup>*Rayapati Venkata Rangarao & Jagarlamudi Chandramouli College of Engineering,  
Guntur, Andhra Pradesh, India*

\* Corresponding author's Email: mohana.kesana@gmail.com

---

**Abstract:** Due to their many applications the optical character recognition (OCR) systems have been developed even for scripts like Telugu. Due to the huge number of symbols utilization, identifying the Telugu words are very much complicated. Pre-computed symbol features have been stored by these types of systems to be recognized or to retrieve in a database. Hence, searching of Telugu script from the database is a challenging task due to the complication in finding the features of the Telugu word images or scripts. Here, we had implemented novel Telugu script recognition and retrieval based on the extraction of texture properties features using iterative partitioned clustering (IPC) for classification of word images. In addition, the statistical feature extraction and similarity matching performance is further improved that measures the similarity between trained and test templates. For testing purpose, we utilized noisy, corrupted and occlusion scanned documents as a query input word images, also considered multi conjunct vowel consonant clustered word images. Our extensive simulation analysis shown that the proposed methodology finds most relevant word images from database even under such conditions. Our proposed scheme has performed superior to the conventional approaches presented in the literature in terms of mean Average Precision (mAP) and mean Average Recall (mAR).

**Keywords:** Telugu script, Texture features, Statistical properties, Iterative partitioned clustering, Statistical parameters, Feature vector, Euclidean distance metric.

---

### 1. Introduction

The procedure of automatic reading of documents is made from a grouping of stages like image acquisition, pre-processing, object extraction, standardization or windowing, feature extraction, arrangement and post-processing. The image acquisition is the way toward assimilating a record as a contribution to the character acknowledgment framework. Pre-processing stage subjects to expulsion of commotions from the image that happens because of assortment of outer components like disgraceful image scan settings, quality or determination of image, quality of image capturing device and absence of enlightenment and so on. The subsequent images require extra pre-processing strategies like disposal of page design [1] and graphical segments in the image [2], skew

recognition and rectification [3] and so on took after by changing the archive to a reasonable shape for additionally processing. The question extraction by and large called as segmentation; the way toward distinguishing limits for region of interest (ROI). The disconnected OCRs guzzle a scanned record information and believers it into machine editable archive design essentially into Unicode of relating character images. The info documents are pre-made with content out of either printed or transcribed content relating to a specific dialect. The documents utilized as a part of character acknowledgment frameworks are named assortment of sorts [4].

The standard arrangements are printed and written by hand documents. The printed documents are created with a specific text style and size of the dialect. The transcribed documents are composed by the writer in a specific content, penmanship style of the person, with the flexibility of composing. The

opportunity and adaptability taken by the writer while composing the content includes more complexities in acknowledgment of the characters, since every individual have their own penmanship style and moreover the errors while composing adds clashes during the time spent acknowledgment. The blend of printed and manually written record had come about into another sort of documents called as application/pre-printed documents or structures [5], comprising of both printed and transcribed contents. Recovery of significant report images to a given inquiry word has developed as a critical theme of research lately. This has developed as an effective other option to the acknowledgment-based recovery for manually written and additionally printed documents in numerous perplexing contents. Practically, these strategies coordinate the question word with an extensive arrangement of words in the database and recover a positioned set of documents considering the similitude (or importance) to the inquiry. Execution of these acknowledgment free recovery plans (regularly called word spotting) relies upon the representation and additionally the coordinating plan utilized for measuring the closeness. In word spotting, word images are represented by fitting highlights. Closeness between words is measured with the assistance of a proper distance work.

Various works are accounted for in writing for characterization or acknowledgment of word images for various dialects. Acknowledgment of words from scanned Telugu documents has not been investigated much when contrasted with the customary scanned documents in English writing. A portion of the methodologies alluded are talked about in a word. Scale-invariant feature transform (SIFT) [6] highlights processed at intrigue focuses are the most well-known highlights for building bag-of-words (BoW) representation. In customary BoW representation, each component point is represented by precisely one "visual word". The word image is represented as a histogram of the visual words. "The quantization of the highlights to a visual word brings about loss of data. Regularly,

this is thought to present some level of power (or invariance). Indeed, even now, there is no agreement on choosing the span of the vocabulary and strategies for taking in the vocabulary. There are two essential strides in developing a BoW representation, given a vocabulary. They are Coding and Pooling. It has been watched that the word reference/vocabulary is over total, and the issue of coding would now be able to be planned as a meagre coding issue with accentuation on learning codes that assistance in depicting the visual substance of the image. In content recovery, each archive is represented by an unordered arrangement of non-unmistakable words present in the record, paying little respect to the language structure and character arrange. Report is formally represented with the assistance of recurrence of events (histogram) of the words in the vocabulary. These histograms are then used to perform record classification and recovery.

In common scene images, spatial request is given by spatial pyramid matching (SPM) [7], which separates images into vertical and flat bearings hence. To give arrange in representation, word images are separated vertically in three sections and afterward recorded. In [8], word images have been represented by profile highlights and Euclidean separation is innocently used to register the comparability. To deal with the fluctuation in word image lengths, word images are coordinated utilizing dynamic time wrapping (DTW) [8]. Nonetheless, our concentration in this work is to infer a recovery conspire with a file structure in the back end. DTW based system isn't reasonable for this reason. Versatility in report recovery has likewise gotten critical consideration in the current past. In [9], 10 Million pages were listed in view of locally likely arrangement hashing (LLAH). In their work, pages are listed, and recovery is done in sub-seconds. This technique concentrates on pages as inquiry and their varieties in imaging, specifically amid the camera catch. Our emphasis is on precise recovery at word level. Late endeavours [10], [11] on strong record recovery utilize visual Bag of Words (BoVW) for representing and coordinating word images. With BoVW representation, and a reversed ordering plan, one can recover important documents from a Million documents in sub-seconds time [10]. Quantization of feature points will be done for the image representation, and a flexible representation is worked by defining a "vocabulary" over an element space. Ideally, vector quantization (VQ) is utilized to produce code from raw descriptors. One restriction of the code-book approach is the hard task of code words in the vocabulary to the feature vectors of an image. The



Figure. 1 Sample images from dataset (Observe different variations in print and noise)

hard task offers two issues: codeword vulnerability and codeword believability. Codeword vulnerability is the issue of choosing the right codeword out of at least two important hopefuls. The VQ approach chooses the best representing visual word, disregarding the importance of different hopefuls. Codeword believability indicates the issue of choosing a codeword without a reasonable hopeful in the vocabulary. The codebook approach allots the best fitting codeword, although this codeword isn't a legitimate representative. To resolve this limitation, authors in [12], [13] presented a soft assignment coding scheme, in which all the visual words will get allotted with a local feature based on their locality. The coding coefficient represents the local features membership function to various visual words. However, the diminishment of reconstruction error is not possible by this approach. The approach of word image retrieval in [30] explains the BoVW scheme with SIFT descriptor in which various languages word images have been utilized for testing the efficiency of SIFT + BoVW with respect to mAP and mAR.

Recently in [24], the author has presented a correlation and hidden markov model (HMM) based Telugu word spotting with superior performance over conventional word image retrieval systems. In [29], author has presented an algorithm based on speeded up robust features (SURF) with the help of BoVW to improve the performance over BoVW and SIFT + BoVW. However, there is a lack of efficiency, higher complexity, lower precision with large data bases are the limitations of the word image retrieval algorithms presented in the literature.

On the other hand, characters in south Indian scripts like Telugu are composed of more than one object making it more complex to apply high level feature extraction techniques. Most importantly, as per author knowledge, Telugu word image retrieval with the noisy, occlusion affected, and random distorted query words hasn't discussed in any paper from the literature.

Therefore, the present work aims at performing the extraction of texture properties features using iterative partitioned clustering (IPC) for classification of word images. In addition, the statistical feature extraction and similarity matching performance is further improved that measures the similarity between trained and test templates.

The novel contributions of this work can be summarized as follows:

- Novel use of extraction of texture properties features using iterative partitioned clustering (IPC) for classification of Telugu word images applications. This approach has not

been used for word image applications before, according to the best knowledge of the authors.

- A fully new framework on Telugu word image retrieval and recognition process is proposed by utilizing a hybrid approach with different sort of distorted word images.
- We mainly aimed at considering the Telugu word images affected with noise, occlusions and random distortions, which was not focused in the past research work done in this field.
- To overcome the drawbacks of BoVW [10], SIFT + BoVW [30], algorithm presented in [24], SURF + BoVW [29] and hash coded hamming (HCH) scheme, our present work aims at performing texture feature extraction given in section 3.1, extraction of properties by computing contrast, correlation, energy and homogeneity (discussed in section 3.2), for classification, we utilized iterative partitioned clustering (discussed in section 3.3). We also calculated the statistical parameters such as mean, variance, standard deviation and entropy followed by similarity measurement for further improvement in the proposed retrieval/recognition system.

The rest of the paper is organized as follows: Section II provides a brief description of the related work done in the field of word image retrieval applications from the past years. The proposed hybrid framework is described in Section III. Section IV explains the results and comparative evaluations against various works presented in the literature. Finally, Section V concludes the paper followed by references.

## 2. Related work

Author in [14] presented that to distinguish a word savvy content; Gradient Local Auto-Correlation (GLAC) feature is extremely hearty and successful and they found that for recognizing or perceiving the Telugu content, the inclination feature is more appropriate and viable than the customary texture features. In [15], three different features named as Gabor, Zernike moments and slope with 400 measurements for word-wise content distinguishing proof and classification has been worked out by utilizing the support vector machine (SVM). These creators have clarified that the essential pre-processing strategies are required to beat the issues with the info or source. A plan named as format coordinating has been used for word acknowledgment in [16] with the feature ID as

Gradient Angular Features (GAF) tried on 760 words from six unique contents. Considering all the exploration papers talked about above and in the writing uncovered that the novel demonstrating of structure of content is a testing undertaking for perceiving or recognizing the word-content. Danish et al. [17] utilized the layout coordinating method for acknowledgment of manually written, machine printed, and sort composed English characters and numerals acquired an exactness of around 94.50% for standard typewritten text styles, 88% for obscure sort composed textual styles, 98% for numerals and 75% if there should be an occurrence of obscure sort composed text styles. Nikhil et al. [18] connected the layout for multi textual style styles and multi text dimensions of English content and accomplished an exactness of around 90%. Mo Wenying et al. [19] connected the layout coordinating calculation by modifying concerning weighted coordinating degree. This calculation gives a higher coordinating rate and conquer the deceptive acknowledgment delivered by conventional figuring strategy with exactness of around 100%. Jatin et al. [20] utilized the format coordinating procedure for sort composed English characters and grouped utilizing neural system classifiers. Soumendu et al. [21] had proposed a calculation for Japanese character acknowledgment utilizing the focal point of gravity features and Euclidean separation features and character with least Euclidean separation is the feature utilized for character acknowledgment. Mahabubar et al [22] had proposed a technique for perceiving Bangla written by hand characters utilizing the convolution neural systems. As a rule, abnormal state feature extraction is utilized for identifying and perceiving the state of a specific protest in the image. Shape depiction shows figuring the position, introduction and size of a specific question in image. The portrayal of a specific shape relies on the application prerequisites. The invariant features of a shape are typically considered to set up the unwavering quality and power of the framework. The articles in image that look for the shape data requires windowing it to a typical size, scale and appropriate brightening to separate the shape from its back-ground force. Numerous option ways likewise existing to play out the coordinating between the test format and prepared layouts in the images [23]. Recently, author in [24] presented a HMM based correlation method for retrieving the Telugu word images. In [25], template matching scheme has been introduced for the retrieval of word images. SURF descriptor-based k-means clustering scheme has been studied and analyzed in [26] where the author considered noisy word inputs for

retrieving the relevant word images from the database. However, lack of efficiency, higher complexity, lower precision with large data bases are the limitations of the word image retrieval algorithms presented in the literature. Additionally, Telugu word image retrieval is a difficult and challenging task due to that each word image has its own structure with single and multi-conjunct vowel consonant cluster. To retrieve those sort of word images, we need to obtain most effective features that describes the relevant data of query word images. This motivates us to implement a high-level feature extraction and a superior clustering algorithm with statistical calculations for Telugu word image retrieval system.

### 3. Proposed frame work

This section describes the proposed model frame work in detail. The block diagram of our proposed system has shown in fig 2, which gives the overall structural procedure of Telugu word recognition and retrieval system. Algorithm 1 describes the overall procedure of proposed hybrid scheme for retrieving the Telugu word images.

---

#### Algorithm 1: Proposed hybrid algorithm

---

**Step 1:** Read and resize the query word image

---

**Step 2:** Convert it into a gray scale image

---

**Step 3:** Now, find the texture features using the gray scale matrix method

---

**Step 4:** Calculate the properties of gray scale texture features such as contrast, correlation, energy and homogeneity

---

**Step 5:** Apply iterative partitioned clustering (IPC) algorithm for segmenting the query word image with 3 clusters

---

**Step 6:** Select the cluster index in which the exact word image has been presented

---

**Step 7:** Now, calculate statistical parameters such as mean ( $\mu$ ), variance (V), standard deviation (STD), root mean square error (RMS) and entropy (E)

---

**Step 8:** Make a query feature vector denoted by 'Fv<sub>1</sub>' by concatenating the obtained features from step 4 and 7

---

**Step 9:** Apply the same process to the database images and obtain the database feature vector denoted as 'Fv<sub>2</sub>'

---

**Step 10:** Now measure the similarity between Fv<sub>1</sub> and Fv<sub>2</sub> using the Euclidean distance(ED) weight function

---

$$ED = \sqrt{\sum (Fv_1 - Fv_2)^2}$$


---

**Step 11:** Finally, display the retrieved Telugu word images, which are most relevant to the query word image

### 3.1 Texture feature extraction

Texture is one of the critical qualities utilized as a part of distinguishing items or districts of enthusiasm for an image. We are worried about the assignment of building up an arrangement of features for ordering or sorting pictorial word information.

The grouping of this information should be possible on a determination cell premise, (for example, in distinguishing the product class of an arrangement cell on satellite imagery) or on a piece of coterminous determination cells, (for example, in recognizing the harvest classification of a whole agrarian field stretching out finished countless cells),

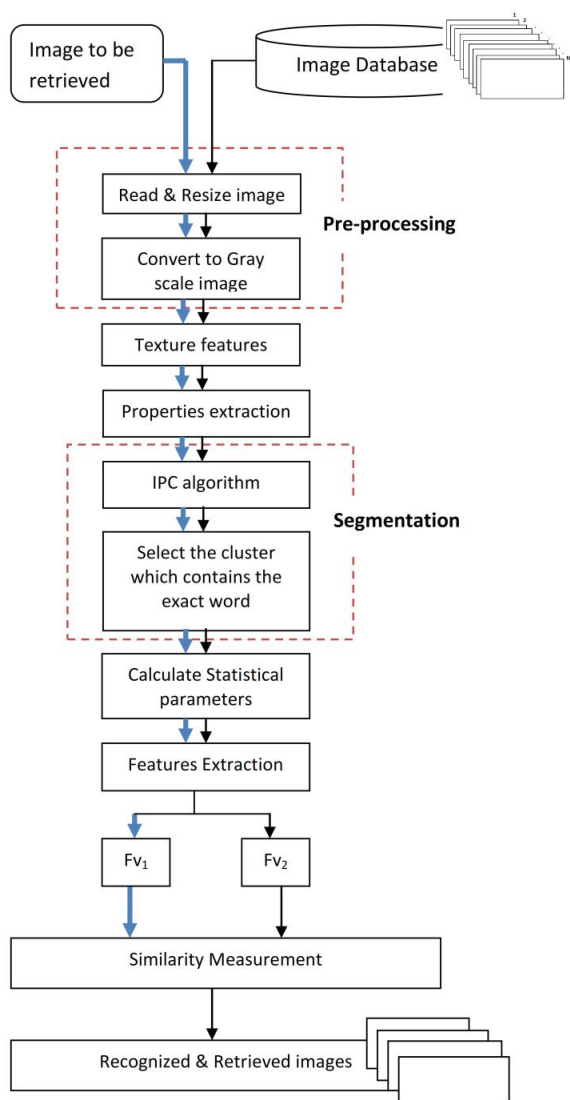


Figure. 2 Block diagram of proposed frame work

the most troublesome advance in ordering pictorial data from a vast square of determination cells is that of characterizing an arrangement of significant features to portray the pictorial data from the piece of determination cells. Once these features are characterized, word image squares can be sorted utilizing anybody of a huge number of example acknowledgment methods. These features are ascertained in the spatial space, and the statistical idea of texture is considered in our technique, which depends on the supposition that the texture data in an image is contained in the by and large or "normal" spatial relationship which the gray tones in the image need to each other. These features contain data about such image textural attributes as homogeneity, gray-tone direct conditions (straight structure), difference, number and nature of limits present, and the multifaceted nature of the image. Note that the quantity of operations required registering any of these features is corresponding to the quantity of determination cells in the image square. It is therefore that we call these features rapidly process able.

### 3.2 Properties extraction

#### 3.2.1. Contrast

It gives the intensity contrast measurement of a pixel to its neighbour over the whole word image

$$Contrast = \sum_{i,j} |i - j|^2 p(i,j) \tag{1}$$

Where  $p$  is a word image with a size of  $i \times j$  in which the number of rows denoted by  $i$  and number of columns denoted by  $j$ .

#### 3.2.2. Correlation

It returns a measure of how correlated a pixel to its neighbour over the whole word image with a range of -1 to 1

$$Correlation = \sum_{i,j} \frac{(i-\mu_i)(j-\mu_j)p(i,j)}{\sigma_i \sigma_j} \tag{2}$$

Where  $p$  is a word image with a size of  $i \times j$  in which the number of rows denoted by  $i$  and number of columns denoted by  $j$ .  $\mu$  represents the mean value and  $\sigma_i, \sigma_j$  are the variances of  $i$  and  $j$ .

#### 3.2.3. Energy

It provides the squared sum of the elements presented in gray scale matrix of word image

$$Energy = \sum_{i,j} p(i,j)^2 \tag{3}$$

Where  $p$  is a word image with a size of  $i \times j$ .

### 3.2.4. Homogeneity

It will be used to measure the closeness of the elements distribution in the gray scale matrix to its diagonal

$$Homogeneity = \sum_{i,j} \frac{p(i,j)}{1+|i-j|} \tag{4}$$

Where  $p$  is a word image with a size of  $i \times j$ .

### 3.3 Iterative partitioned clustering (IPC)

Here in the proposed clustering method, we treat that every object as having a location in space. Overall structure of IPC algorithm utilized in this scenario is depicted in Fig. 3. Internal programming of IPC algorithm is demonstrated in algorithm 2, where it utilizes a word image as a dataset denoted by  $I$ , number of clusters is represented as  $K$ , distance metric as distance and  $x, y, rep$  and  $eps$  are known to be global variables. The functions utilized in this algorithm like  $distfun$ ,  $accumarray$ ,  $find$ ,  $Dist$ ,  $sum$ ,  $min$  and  $reshape$  are pre-defined MATLAB functions. The  $distfun$  computes the distance between every point from  $I$  to each cluster in  $K$ .

#### Algorithm 2: IPC algorithm

$I$ : input data set

$K$ : number of clusters

Distance: Euclidean distance

Global variable:  $x, y, rep, reps$

Calculate the distance between each cluster to input data set and the primary allocation of points to cluster centroids

$$Dist = distfun(I, K, distance, 0);$$

$$[ds, indx] = min(Dist, [], 2);$$

$$x = accumarray(indx, 1, [z, 1]);$$

Compute the distances sum based on cluster-wise priority

$$nonempties1 = find(x > 0);$$

$Dist(:, nonempties1)$

$$= distfun(I, K(nonempties1, :), distance, iter);$$

$$ds = Dist((indx - 1) * y + (1: y)');$$

$$sumDist = accumarray(indx, ds, [z, 1]);$$

$totsumDist$

$$= sum(sumDist(nonempties1));$$

```

iter = iter + 1;
reshape;
End
SegmentedOutput;
End
    
```

Function  $min$  finds the indices of the minimum values and returns them in an output vector.  $accumarray$  returns an array by accumulating of elements. All the indices and non-zero elements will be returned by the function  $find$ . Function  $sum$  utilized to calculate the sum of elements along the first array dimension.

### 3.4. Statistical Parameters

Statistics is the investigation of the gathering, association, examination, and interpretation of information. It manages all parts of this, including

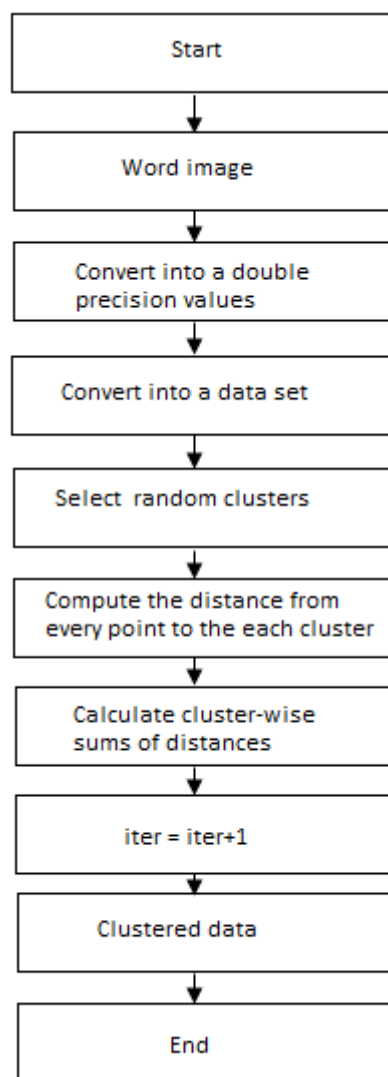


Figure. 3 IPC algorithm flow chart

the arranging of information gathering as far as the plan of studies and tests. The different statistical measures [27] are mean ( $\mu$ ), difference (V), standard deviations (STD), root mean square error (RMS) and entropy (E). These measures are utilized as a part of an extensive variety of logical and social research, including: biostatistics, computational science, computational human science, organize science, sociology, humanism and social research and so forth. In this, we have used different sort of statistical measure regarding image processing and mimicked these.

**3.4.1. Mean ( $\mu$ )**

$$Mean (\mu) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} I(x, y) \quad (5)$$

Where I = input word image  
 x and y = spatial co-ordinates  
 M and N are number of rows and columns

**3.4.2. Variance (V)**

It is a measure of how far an arrangement of numbers is spread out. It is one of a few descriptors of a likelihood appropriation, portraying how far the numbers lie from the mean (expected esteem). Specifically, the difference is one of the snapshots of a circulation. In that specific circumstance, it shapes some portion of a deliberate way to deal with recognizing likelihood disseminations. While other such methodologies have been created, those in view of minutes are favourable regarding scientific and computational straightforwardness. Scientifically fluctuation is given by

$$Variance (V) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \left( I(x, y) - \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} I(x, y) \right)^2 \quad (6)$$

Where I is an input word image  
 x and y are spatial co-ordinates  
 M and N are number of rows and columns

**3.4.3. Standard deviation (STD)**

It is a most broadly utilized measure of changeability or assorted variety utilized as a part of statistics. As far as image processing it indicates how much variety or "scattering" exists from the mean or expected esteem. A low standard deviation shows that the information guides incline toward be near the mean, while exclusive requirement deviation demonstrates that the information calls attention to spread out finished an expansive scope of qualities.

**STD**

$$= \sqrt{\frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \left( I(x, y) - \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} I(x, y) \right)^2} \quad (7)$$

**3.4.4. Entropy (E)**

Entropy is a statistical measure of randomness that can be used to characterize the texture of the input image.

$$Entropy(E) = - \sum_i H_i \times \log_2 H_i \quad (8)$$

Entropy is defined as

- Calculate histogram counts ‘H’
- Remove zero entries in H
- Normalize H so that sum (H) is one

**4. Results and discussion**

The experimental results and analysis is obtained by MATLAB tool to show the performance of proposed system. We have used scanned documents of Telugu language for evaluation purpose. Details are provided in the Table 1 and sample word images are shown in Fig. 1. All the books are annotated at the word level and ground truth was created using technique described in [28]. To evaluate the quantitative performance, multiple query images were generated. The query images are selected such that (i) They have multiple occurrences in the database, (ii) They are mostly functional words and (iii) They have no stop words. The performances of proposed method are measured by mean Average Precision (mAP) and mean Average Recall (mAR) and the obtained results are compared with the BoVW [10], SIFT+BoVW [30], algorithm presented in [24] i.e., HMM with correlation [24]SURF + BoVW [29] and HCH scheme.

For testing purpose, we considered the Telugu word images with occlusions, Gaussian noisy and corrupted or sparse as a query word image. We got retrieved the relevant images with the effected word images as a query. This helps us to recognize or retrieve the Telugu word images even after corrupted by human or any other natural defects. The simulated outputs of retrieved word images with a query word image of ‘pariksha’ has been

Table 1. Book used for experiment

Book	#Pages	#Words
Telugu	425	7425

shown in Fig. 4, we can observe that there is a noise in the input query but still the retrieved word images are more relevant to the input query, which shows the robustness and effectiveness of proposed method where this is not possible by the algorithms presented in the literature [10, 24, 29, 30]. Figs. 5 and 7 shown the retrieved output images of multi-conjunct vowel consonant clustered word images with the effects of occlusion and corruption. The obtained quality metric values Like mAP and mAR are presented in Table 2, where BoVW presented in [10] has obtained the values of mAP is 81.7% and mAR is 77%, algorithm given in [30] got the values

of mAP and mAR are 85.3% and 79.9 % respectively. Reference [24] presented a betterment with HMM based correlation by getting the enhanced mAP and mAR as given in Table 2. HCH scheme has been proposed by the author itself and is accepted for publication in springer lecture notes in electrical engineering (LNEE) and our proposed system obtained the values of mAP is 96.7% and mAR is 84.2%, which are much superior to the results presented in the Table 2. Performance evaluation of conventional and proposed Telugu word retrieval system is disclosed in Fig. 8.

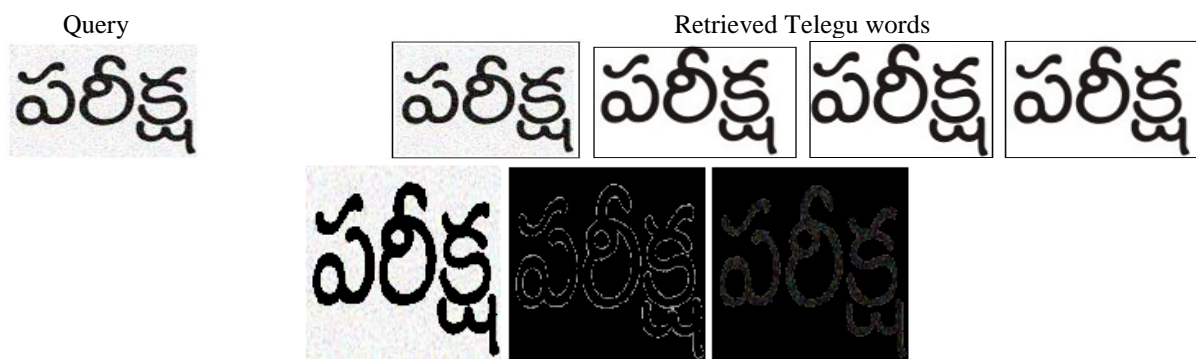


Figure. 4 Noisy query word image, retrieved word images from DB and segmented output of IPC algorithm



Figure. 5 Corrupted query word, retrieved word images and segmented output of IPC algorithm Note that the irrelevant word images shown in red color



Figure. 6 Corrupted query word, retrieved words from DB and segmented o/p of IPC algorithm





Figure. 7 Query word with occlusion, retrieved words from DB and segmented output obtained by IPC algorithm

Table 2. mAP and mAR of different schemes on different books in dataset

Measurement	BoVW [12]	SIFT + BoVW [30]	Ref [24]	Ref [29]	HCH scheme	Proposed
mAP	0.817	0.853	0.89	0.731	0.91	<b>0.967</b>
mAR	0.77	0.799	0.823	0.809	0.835	<b>0.842</b>

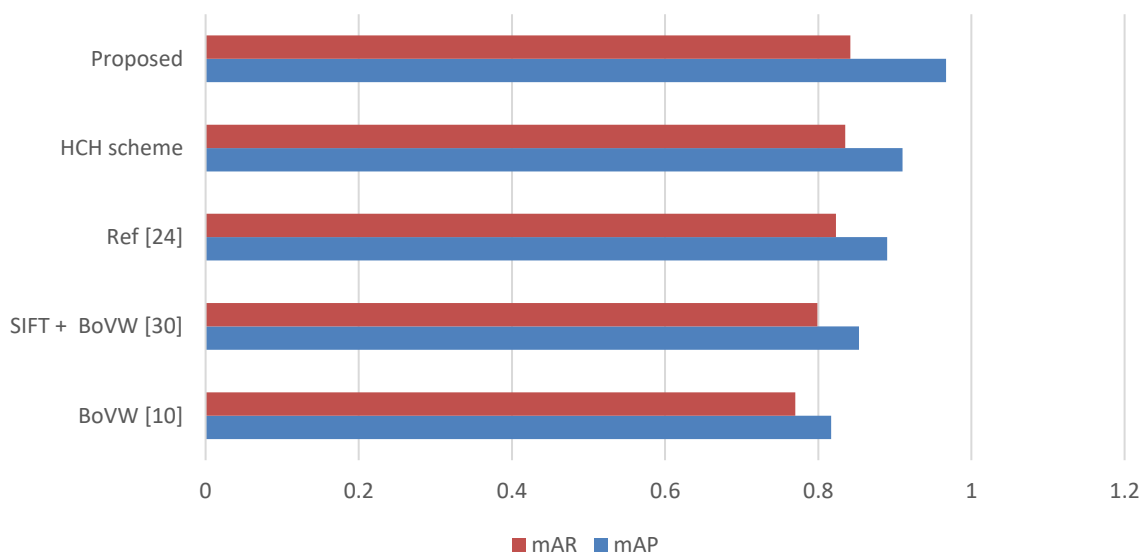


Figure. 8 Performance evaluation of conventional and proposed word retrieval methods

### 5. Conclusions

In this work, we implemented a new framework for Telugu word image retrieval and recognition system by utilizing a hybrid approach with different sort of distorted word images as a query. We mainly focused to get the most relevant images even the query word image gets affected with noise, occlusions and random distortions. Our present work aims at performing texture feature extraction by computing contrast, correlation, energy and homogeneity and for classification, we utilized IPC algorithm. We also computed the statistical parameters such as mean, variance, standard deviation and entropy followed by similarity measurement for further improvement in the

proposed retrieval system. Extensive simulation analysis disclosed that the proposed hybrid methodology achieved superior performance over the conventional retrieval system by achieving higher mAP and mAR.

Our future enhancements include that the maximum number of features will be extracted to obtain most possible relevant images with various types of noisy query word images from Telugu literature by introducing some novel contributions to achieve higher mAP and mAR with meta data bases.

### References

[1] B. Verma, M. Blumenstein, and S. Kulkarni, "Recent achievements in off-line handwriting recognition systems", *School of Information*

Technology, Griffith University, Gold Coast Campus.

- [2] C. V. Jawahar, M. N. S. S. K. P. Kumar, and S. S. R. Kiran, "A Bilingual OCR for Hindi-Telugu Documents and its Applications", *Centre for Visual Information Technology, International Institute of Information Technology, Hyderabad*.
- [3] N. S. Rani, and T. Vasudev, "A Generic Line Elimination Methodology using Circular Masks for Printed and Handwritten Document Images", *Emerging research in computing, information, communication and applications*, Vol. 3, No. 1, pp. 589-594, 2014.
- [4] R. Singh and M. Kaur, "OCR for Telugu Script Using Back-Propagation Based Classifier", *International Journal of Information Technology and Knowledge Management*, Vol. 2, No. 2, pp. 639-643, 2010.
- [5] S. V. Patgar, T. Vasudev, and S. Murali, "A system for detection of fabrication in photocopy document", *Journal of Computer Science & Information Technology*, Vol. 5, No. 14, pp. 29-35, 2015.
- [6] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints", *International Journal of Computer Vision*, Vol. 60, No. 2, pp. 91-110, 2004.
- [7] S. Lazebnik, C. Schmid, and J. Ponce, "Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories", In: *Proc. of IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, 2006.
- [8] T. M. Rath and R. Manmatha, "Word spotting for historical documents", *International Journal of Document Analysis and Research*, Vol. 9, No. 2-4, pp.139-152, 2007.
- [9] K. Takeda, K. Kise, and M. Iwamura, "Real-time document image retrieval for a 10 Million pages database with a memory efficient and stability improved LLAH", In: *Proc. of the International Conf. on Document Analysis and Recognition*, pp. 1054-1058, 2011.
- [10] R. Shekhar and C. V. Jawahar, "Word Image Retrieval Using Bag of Visual Words", In: *Proc. of the International Workshop on Document Analysis Systems*, pp. 297-301, 2012.
- [11] I. Z. Yalniz and R. Manmatha, "An Efficient Framework for Searching Text in Noisy Document Images", In: *Proc. of the International Workshop on Document Analysis Systems*, pp. 48-52, 2012.
- [12] J. van Gemert, J.-M. Geusebroek, C. J. Veenman, and A. W. M. Smeulders, "Kernel Codebooks for Scene Categorization", In: *Proc. of European Conf. on Computer Vision*, pp. 696-709, 2008.
- [13] J. van Gemert, C. J. Veenman, A. W. M. Smeulders, and J.-M. Geusebroek, "Visual Word Ambiguity", *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 32, No. 7, pp.1271-1283, 2010.
- [14] N. Sharma, S. Chanda, U. Pal, and M. Blumenstein, "Word-wise Script Identification from Video Frames", In: *Proc. of International Conf. on Document Analysis and Recognition*, pp.867-871, 2013.
- [15] N. Sharma, U. Pal, and M. Blumenstein, "A Study on Word Level Multi-script Identification from Video Frames", In: *Proc. of International Joint Conf. on Neural Networks*, pp.1827-1833, 2014.
- [16] P. Shivakumara, N. Sharma, U. Pal, M. Blumenstein, and C. L. Tan, "Gradient-Angular-Features for Word-wise Video Script Identification", In: *Proc. of International Conf. on Pattern Recognition*, pp.3098-3103, 2014.
- [17] D. Nadeem and S. Rizvi, *Character recognition using template matching*, Department of Computer Science, Jamia MilliaIslamia, New Delhi, 2015.
- [18] N. R. Pai and V. S. Kolkure, "Design and implementation of optical character recognition using template matching for multi fonts size", *International Journal of Research in Engineering and Technology*, Vol. 4, No. 2, pp. 398-400, 2015.
- [19] M. Wenying and D. Zuchun, "A Digital Character Recognition Algorithm Based on the Template Weighted Match Degree", *International Journal of Smart Home*, Vol. 7, No. 3, pp. 53-60, 2013.
- [20] J. M. Patil and A. P. Mane, "Multi Font and Size Optical Character Recognition Using Template Matching", *International Journal of Emerging Technology and Advanced Engineering*, Vol. 3, No. 1, pp. 504-506, 2013.
- [21] S. Das and S. Banerjee, "An Algorithm for Japanese Character Recognition", *International Journal of Image, Graphics and Signal Processing*, Vol. 7, No. 1, pp. 9-15, 2014.
- [22] M. M. Rahman, M. A. H. Akhand, S. Islam, P. C. Shill and M. M. H. Rahman, "Bangla Handwritten Character Recognition using Convolutional Neural Network", *International Journal of Image, Graphics and Signal Processing*, Vol. 7, No. 8, pp. 42-49, 2015.
- [23] J. S. Kiran, N. V. Kumar, N. S. Prabha, and M. Kavva, "A Literature Survey on Digital Image Processing technique in character recognition

- of Indian languages”, *International Journal of Computer Science and Information Technologies*, Vol. 6, No. 3, pp. 2065-2069, 2015.
- [24] D. Nagasudha and Y. M. Latha, “Keyword Spotting using HMM in Printed Telugu Documents”, In: *Proc. of International Conf. on Signal Processing, Communication, Power and Embedded Systems*, pp. 1997-2000, 2016.
- [25] T. N. S. R. Vasudev and C.H. Pradeep, “A Performance Efficient Technique for Recognition of Telugu Script Using Template Matching”, *International Journal of Image, Graphics and Signal Processing*, Vol. 8, No. 3, pp.15-23, 2016.
- [26] K. M. Lakshmi and T. R. Babu, “Searching for Telugu Script in Noisy Images using SURF Descriptors”, In: *Proc. of IEEE 6<sup>th</sup> International Conference on Advance Computing*, pp. 480-483, 2016.
- [27] E. Kreyszig, *Advanced Engineering Mathematics*, J. Willey & Sons Inc., 2011.
- [28] C. V. Jawahar and A. Kumar, “Content-level Annotation of Large Collection of Printed Document Images”, In: *Proc. of International Conf. on Document Analysis and Recognition*, 2007.
- [29] N. Jayanthi and S. Indu, “Inscription Images Retrieval Using Bag-of-Visual Words”, *IOP Conference Series: Material Science and Engineering*, Vol. 225, No. 1, pp. 1-8, 2017.
- [30] R. Shekhar and C. V. Jawahar, “Word Image Retrieval Using Bag-of-Visual Words”, In: *Proc. of IAPR International Workshop on Document Analysis Systems*, pp. 297-301, 2012.