An Efficient Telugu Word Image Retrieval System using Deep Cluster

Kesana Mohana Lakshmi[a], Tummala Ranga Babu[b]

[a] Research Scholar ,Dept of ECE,University college of engineering & Technology,Acharya nagarjuna university, Guntur
Associate Professor,Dept of ECE,CMR Technical Campus, Hyderabad, Telangana, India
[b] RVR & JC College of Engineering, Guntur, Andhra Pradesh, India
[a] mohana.kesana@gmail.com, [b] trbaburvr@gmail.com

Article History: Received: 11 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 10 May 2021

Abstract: Optical character recognition (OCR) system has successfully implemented in many applications of text classifications and text retrieval systems. But, due to the complexity of huge number of alphabets (symbols), grammatical punctuations,ithas failed to classify and retrieve many regional languages. Thus, to overcome these problems, word image retrieval is innovated. this article presents the Telugu word image retrieval(TWIR) system using advanced artificial intelligence based multi-layer deep learning convolution neural network (DL-CNN) is developed for extraction of accurate Telugu word features. Then, AlexNet based deep learning model is used for further clustering of features and the procedure named as DeepCluster. Thus, by using DeepCluster method logy various grammar rules of Telugu scripts will be perfectly analyzed. Hence, the feature database is trained with the Telugu word features along with its grammar rules. Thus, the system will effectively use for retrieval Telugu words and useful in all real time applications such as English to Telugu conversions and Telugu web browsing, Telugu speech analysis. The simulation results show that the proposed deep learning model gives the outstanding results compared to the state of art approaches for both Average Precision (mAP) and mean Average Recall (mAR) metrics.

Keywords: Telugu word image retrieval, deep learning, convolutional neural networks, feature extraction, precision and recall

1. Introduction

Image is an essential part of daily life[1]. There are many images of data generated every day by using electronic gadgets. Image retrieval[2] has been broadly classified into two type namely text-based approach and content-based approach. Each approach has its own unique features[3] that can be used in various application based their scenario.Content based image retrieval (CBIR) [4] also known as query by image content. With this approach that the search analyses the visual content of the image rather than keyword descriptions associated with the image [5]. Text based image retrieval (TBIR) [6]purely based on text keywords descriptions used as input and Index images also used text keywords. By comparing both input keywords and index images keywords, the matching images retrieved from image repository. Advantages of this approach are easy to implement, fast retrieval and Web image search. Manually searching for any particular image from a large set of images is not an easy task. TBIR is a technique that is used to search an image in a database of the image using text annotation. But TBIR is a very time taking and not an efficient technique because in the TBIR technique user has to check every image one by one manually. Sometimes different users have different perceptions about the images. For example, if one image containing greenery and animals, both then users can name that image based on greenery or based on the animal's name. Disadvantages are annotation of images manually is not available at most of the time; annotation is difficult for a large image repository when manually annotated and manual annotation is not accurate. In TWIR, there are two challenges such as intention gap and semantic gap [7]. The first one refers to expected visual content or image not exactly retrieved by the user through a query and second one refers to difficulty in relating high-level semantic concept with low-level visual feature. In order to overcome these two challenges, researchers have focused their efforts towards the TWIR systems[8].With this approach visual content plays a major role in retrieval of word images such as texture, color, intensity, shape, resolution etc.[9]. At present scenario image retrieval domain has faced lots of challenges [10] due to its growth of digitization in day-to-day activities. Though there are diverse kinds of word grammar issues in the script domain, symbol dentistry issues are one of the major problems. To manage such symbol issues,proposed TWIR informaticsrepositories play an important role.

The major contribution of the proposed methodology as follows:

- Initially all the images are normalized to the fixed size, so all the train and test image features will be perfectly estimated. This Normalization process also used to calculate and analyze the various distortions such as missing segment with distortion, random distortion, noisy effected, missing segment in Query words.
- The DLCNN approach is applied to extract both textural and statically features from Telugu word. The major significance of DLCNN is to analyze word image and analysis it with respect to Telugu symbols.
- Then, DeepCluster process is applied on the feature vectors by using the AlexNet deep learning model. Thus, here input features will be divided into various classes based on their grammatical significance and rules.

- Finally, pairwisehamming distance measurement is used to compare the various features with database along with its classes and retrieves the similar resultant output images based on their indexes.
- The proposed results are compared to the various litterers such as SDM-NSCT [17], HWNET v2 [16], GLCM-IPC [15], SURF-BoVW [14], HMM-C [13], and SIFT-BoVW [12] and provide the better Precision and Recall accuracy.

Rest of the paper organized as follows; Section 2 gives the detailed expiation of various states of art approaches with their drawbacks. Section 3 gives the detailed procedure of proposed TWIR system using training and testing phases with DLCNN based feature extraction and DeepCluster based feature clustering. Section 4 gives the detailed analysis of the simulation results of proposed approach and compared with the existing methods. Finally, Section 5 concludes the article with possible future enhancements.

2. Related Work

In [11] authors proposed the image retrieval system using LBP feature extraction. LBP used the local content of any image. There are three types of content of an image, namely color content, shape content, and texture content. LBP focuses on the texture content of an image. LBP technique considers the eight local neighbors of every pixel and calculates the difference between the intensity of the central pixel and neighborhood pixels and generates an eight bits binary pattern. For the corner pixels, the intensity of only three pixels is available. For the rest of the five pixels, padding is used, and we padded intensity with the zero values. In [12] authors applied Bag of visual word (BovW) features for word image retrieval,by refining classification accuracy using fused SIFT and BovW textural features. This work outperforms for normal word images classification by using fused LBP and BovW features. This feature gave more accurate result when compared to other image features. But again, there are more features are available to enhancing the classification accuracy for effective image retrieval system.In [13] authors proposed textural feature extraction based on Hidden Markova Model(HMM) on the printed Telugu documents.To improve the performance they used HMM based coarseness textural feature. The feature extraction of the HMM model is degraded due to it does not support the statistical features, hence reduced in retrieval accuracy.

In [14] authors proposed a retrieval system using SURF feature extraction along with theBoVW model.With this paper authors have studied use of SURF+BoVW model to improve the digital word images for classifying various problems in script readings. SURF is more suitable for extracting image features like bright or dark in both back grounds as well as foregrounds of the image. With this justification the work has been developed and results also improved in a certain amount. The challenges of classifying word images are still motivated among researchers and academicians due to that size of images. But this method results in low precision values.In [15] authors proposedthe TWIR system using iterative portion clustering (IPC) of Telugu word data and Grey level co-occurrence matrix (GLCM) based feature extraction it results in reduced retrieval efficiency. In [16] authors proposed word image retrieval by texture characterization byHWNET v2 based deep learning system. But this deep learning model has the high computational complexity results in more training and testing time compared to other approaches.In [17] authors implemented the TWIR system using SDM based NSCT feature extraction.Here deep convolution neural network used for classifying Telugu word images and during training, learned features and the classification results are used to retrieve word images.But this system gives the less accuracy towards the various distortions and noises.In [18] authors proposed a heterogeneity-aware multi-resolution Local Binary Pattern (hmLBP) for retrieval of histopathology images. Initially in hmLBP, texture features were extracted from different resolution of histology image and texture features were combined which acted as final feature vector. Then, rotation invariant binary codes were used to extract compact binary codes. It reduced dimensionality of feature vector and comprised the vast majority of texture patterns. Finally, the histogram bins were weighted through counted LBP codes. However, it has low F1 measure.

In [19] authors presented multiple classifier color image retrieval using region growing method. They proposed synthetic features to describe interior structure and global shape for the trademark image retrieval. They combined shape description and feature matching to retrieve trademark images.In [20] authors presented an integration of global-local descriptors to extract the features through Zernike's moment's coefficients and edge gradient co-occurrence matrix. The regional and boundary feature fusion for trademark image retrieval. To improve the performance of TIR, we have used entropy as refinement layer to remove the dissimilar trademark images and then applied Zernike moments and SURF feature descriptor to retrieve most similar images.

3. Proposed Method

The main idea of this work is to benefit from the powerful performance of DL-CNN aiming to extract effective features while minimizing time and resources. CNN is the convolutional phase that works like a visual descriptor to extract features from images. Each image undergoes a transformation through an application of a set of filters creating new forms of images called convolution maps. These convolution maps are concatenated into a feature vector called CNN code. All features are saved as an index of features. This process has trainingphase that is intended to extract the features from the dataset's images and atest phase that is used to extract features from the query. Our approach takes place in two phases and is represented in Figure 1:

For retrieval process, similarity comparison technique has been used between test user query image and training classified image feature database. After comparison, resulting images are identified and retrieved based on their shortest distance. The shortest distance images between query image and classified image feature database will be considered as first ranked image for the best match retrieval.

- **Training phase:** In this phase, the features are extracted from the dataset using DL-CNN, and then component analysis based DeepCluster clustering algorithm is used to cluster images into similar groups and determines the gravity center of each group. The gravity centers construct the output layer in the retrieving phase and will be used as an identifier for each group.

- **Test phase:** In this phase, the similarity between features extracted from a query using DLCNN, and the output layer of a CNN to predict the identifier closer to it is calculated. Based on that calculation, the closest cluster is returned. Then, when the similarity is calculated between the query and the images in that cluster, the top-ranking images are finally retrieved.

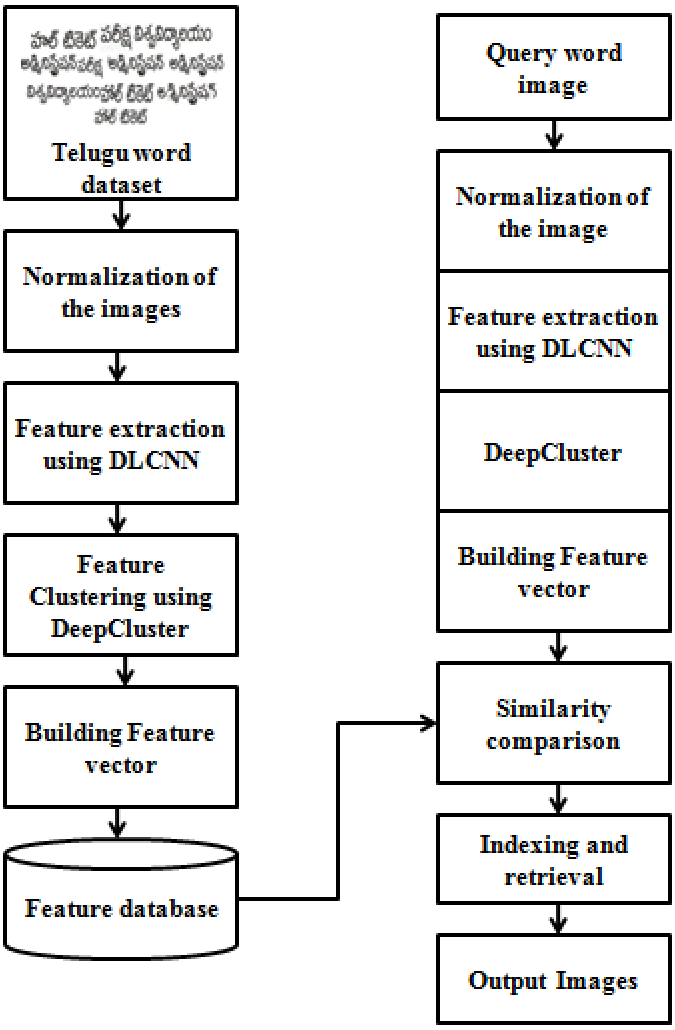


Figure 1: Proposed TWIR model

3.1. Normalization of the Images

For image normalization, all the random sized images are resized into same sized images to extract multiple features of the image on the same structure. Here random sized images are normalized as 200x400 size pixel images without affecting the aspect ratio of the image using bilinear standard transformation. This Normalization process also used to calculate and analyze the various distortions such as missing segment with distortion, random distortion, noisy effected, missing segment in Query words.

3.2. Feature Extraction using DLCNN.

For Feature extraction, the most popular deep learning algorithm so called DLCNN has been used. A CNN is a neural network process which changes a function into some other form to get more details of the image by performing the element wise matrix multiplication operation between original image and filter or kernel image. Here DLCNN used for extracting Telugu word image features such as shape and texture since Telugu word images are in different shape and texture format.

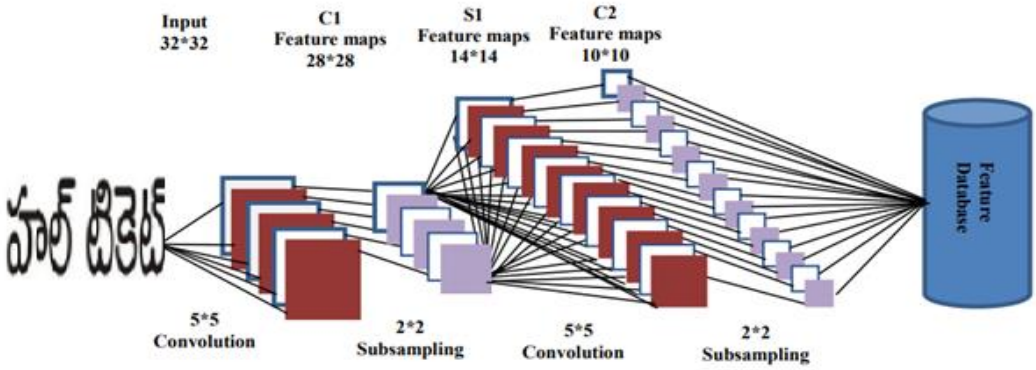


Fig.2: The process of Convolution for feature extraction.

Figure 2 shows the proposed DLCNN model for extraction of features. In convolution operation there are two matrixes in which one is image matrix and another one is filter or kernel matrix that transforms image into another format. The transformed images are called convoluted image. In this work normalized Telugu word image matrixes are given as input to the convolution neural network to convert convoluted image matrix from where features are extracted to build feature vector for the retrieval process. The sample Telugu word input image matrix and corresponding convoluted image matrix using CNN filter or kernel. As we have used 5X5 matrixes for input image, for this example total pixels points are 25.After applying convolution computation that has been reduced as 9-point pixels since that convolution filter or kernel has 3X3.This process has been applied until that fine details of features extracted.

The entire process of convolution process starting from raw Telugu word images into building feature vector for the retrieval process. Here we have used $32 \times 32 \times 1$ image for entire convolution process. Firstly, 5 x 5 kernels with one stride are used over a $32 \times 32 \times 1$ image as convolution layer and $28 \times 28 \times 6$ output matrix is generated. The feature map is reduced as 28×28 from 32×32 with the one stride and no padding. Secondly average pooling method has been used for dimension reduction. For dimension reduction, we have used a filter size of 2×2 with a stride of 2 and dimension was reduced by the 2 factor. This would yield a $14 \times 14 \times 6$. Likewise, another convolution layer was used with sixteen 5×5 kernels to get an output matrix of$10 \times 10 \times 6$. Subsequently another pooling layer is used that yields an output matrix of $5 \times 5 \times 16$.Finally, sixteen 5×5 feature maps extracted from each image and each feature maps consist of width of 5×5 matrixes.In this work, the fully connected layers in the pre-trained DLCNN models are excluded from the network. That means, only the convolutional layers remained in the model for features extraction. The advantage of using those pre-trained images is that they have learned to extract features from a diversity of images.The signature of each Telugu word image was represented by a vector that includes features extracted from DLCNN model. Then, based on those signatures, we cluster the images.

3.3Feature clustering

In this section, the proposed DeepCluster is proposed for cluttering of features. The convolution layers and pooling layers extract the textual features and image features. Then extracted features are processed in the clustering layer which is introduced in the fully connected layer. It takes advantage of high capacity of an optimized AlexNet and the ability of learning the clusters without labels. The cluster labels are used in the fully connected layer to classify the images as relevant and non-relevant images. The relevant images are ranked based on the minimum distance between relevant images and the database images. Figure 3 shows the overall flow of this DeepCluster method.

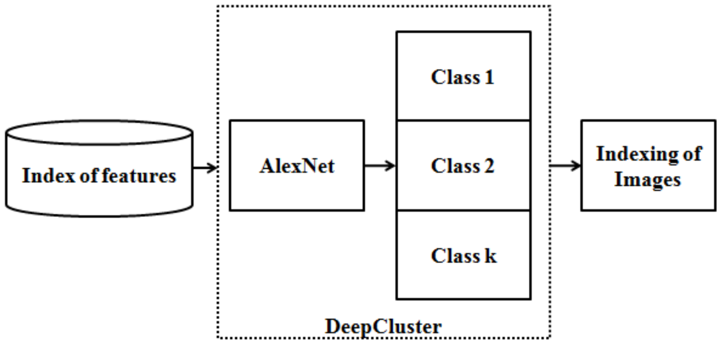


Figure 3: DeepCluster approach.

The textual description and the collected images are given as input to optimized AlexNet. The convolution layer and pooling layer are processed to extract the textual features and visual features. In the classification task, an optimized AlexNet aims to generate predictions from labeled dataset. By learning this mapping using optimized AlexNet, the dimension of the data space is reduced to a much smaller space with the help of max-ave pooling layers and fully convolutional layers. The optimized AlexNet summarizes the features of textual information of the images and visual images. As a result, each layer of optimized AlexNet can be interpreted as a feature embedding of the data.

Generally, a clustering algorithm groups an unlabeled dataset into clusters. The clustering algorithm has the benefit of being able to generate clusters directly from the data without any labels. But, uninformed of class labeling task and are adversely affected by the curse of dimensionality. The clustering algorithm performs better as the dimension of the data space is reduced. The main intention of DeepCluster is to combine the optimized AlexNet and clustering process for efficient TWIR. In the deep embedded clustering, optimized AlexNet learns a mapping between the textual information of images and visual images in the dataset to the query. The embedded clustering technique groups the images into different categories from lower-dimensional embedding. The optimized AlexNet consists of five convolution layers, three pooling layers and three fully connected layers. The convolutional and pooling layer is used to extract the textual features and visual features of images. The DeepCluster is designed such that the second fully connected layer of optimized AlexNet acts as an embedded clustering layer. It is a fully connected layer without Maxout activation, with number of units equal to the number of classes of the dataset. The embedded clustering layer uses a t-distribution kernel to measure the similarity between embedding textual and visual features of images and mean of a cluster distribution.

The probability of an image to belong to a cluster is the embedding textual and visual features of images and is the hyper-parameter. The first fully connected convolutional layer extracts the global features of textual visual information of images. The second fully connected layer of optimized AlexNet learns the centroids of different clusters representing each class. This layer is used to learn a mapping between the embedding textual and visual features of images and the predictions. Moreover, the frequencies are designed to prevent distortion from large clusters. A Kullback-Leiblerdivergence is computed as loss function to compare target image and embedding probability distributions. For the clustering layer, the loss associated with the cluster centers and the embedding of textual and visual features of images.The cluster center (i.e., centroids) with minimum loss is used for further process.

3.3.1 Indexing of Images

The centroids of each cluster are used as class label which is given as input to the next layer. It maps the cluster labels with the relevant images by using activation function. the sigmoid logistic regression function which controls the output between is the cluster labels obtained from the clustering layer, are the weights in this layer and are the bias in the last fully connected layer. Then the binary operation is performed where the feature vectors of the relevant images are mapped into binary codes. After the classification of images, relevant images are ranked based on the Pairwise hamming distance of binary codes between relevant images and database images.

3.4. Retrieving images using DLCNN

The retrieving phase is test and requires a Pairwise hamming distance measure to search for similar images. First, the DLCNN node (already stored as small signature files) from the query is compared to the output layer of the convolution phase. The comparison is taken based on their Pairwise hamming similarity scores between the features. The closest centroid to the query represents the identifier of the relevant class. Finally, we retrieve the top images from the selected cluster. The relevant image which has minimum Pairwise hamming distance has the topmost rank.

Table 1: Proposed TWIR procedure.

Step 1: Load the image from the database.
Step 2: Extract local features of both textual and visual information of image by processing convolutional layer and max-pooling layer.
Step 3: Extract global features of both textual and visual information of images by processing first fully connected layer.
Step 4: Cluster the features by processing the clustering layer
Step 5: Classify the images as relevant and non-relevant images based on the cluster labels by processing the hidden layer.
Step 6: Map the relevant image with the binary codes by performing binary operation.
Step 7: Calculate the Pairwise hamming distance between the relevant images and database images and rank the images based on the Pairwise hamming distance.

4. Results And Discussion

This section gives the detailed analysis of simulation results of proposed system implemented using MATLAB 2019a programming environment. To effectively implement the proposed system, standard books are considered from various Indian digital libraries such as Digital library of India and The universal library respectively. The proposed system trained with the various types of books obtained from these libraries. Table 2 presents the various types of Telugu books such as Novel, arts, comers, history, philosophy, psychology, engineering, architecture, mythology, and so on considered for training process with the available pages and words. Nearly overall one million words are used for the training procedure.

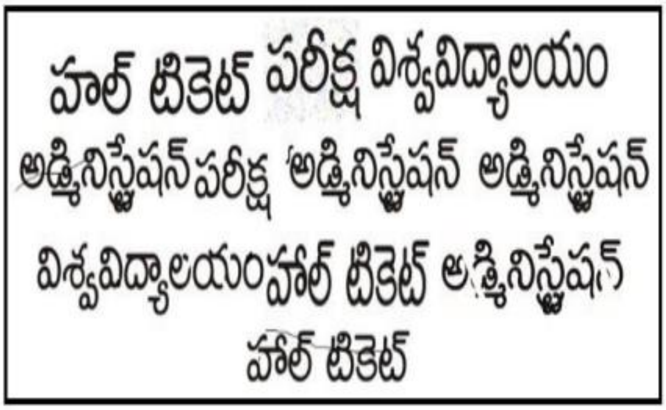


Fig. 4: Test images utilized for TWIR system.

Table 2: Book used for experiment.

Telugu Book Type	No. of Books	Averagepages per each book	Total words
History and Mythology	2	200	22048
Arts and Commerce	3	350	32038
Telugu Literature	7	280	37501
Science and Technology	7	300	64993

Figure 4 gives the sample test Telugu word images with various Grammaticus and distortions. The distortions are missing segment with distortion, random distortion, noisy effected, missing segment. These distortions are generated due to various noise sources, typing mistakes caused by humans, or due to machine printing and scanning problems. Thus it is necessary to analyze the performance of proposed system using precision and recall metrics. For analyzing the performance of proposed TWIR system, mAR and mAP metrics are calculated and compared with the state of art approaches. They are SDM-NSCT [17], HWNET v2 [16], GLCM-IPC [15], SURF-BoVW [14], HMM-C [13], and SIFT-BoVW [12]. The proposed simulation are performed by using various types of test images and gives the enhancive mAP and mAR, because proposed method utilizes the advanced deep learning architectures for feature extraction and feature clustering based on the grammatical rules.

Figure 5 represents the retrieved output Telugu word images for various query words with distortions. Column two represents the all-query word images and whereas column three represents the retrieved word images to that of query word. Figure 5a, 5e, 5f, and 5g represents the scenario of retrieved images for Missing segment with random distortion and Missing segment words-based query images. Even though some parts of Telugu symbols are missed, by using Deep feature clustering output images are retrieved very precisely.Figure 5d represents the scenario of retrieved images for Random disturbance-based query images. Even though lots of disturbances occurred in both trained and testing images, by using DLCNN feature learning with normalization output images are retrieved very precisely. Figure 5c represents the scenario of retrieved images for Random disturbance-based query images. Even though lots of disturbances occurred in both trained and testing images, by using DLCNN feature learning with normalization output images are retrieved very precisely.Figure 5c represents the scenario of retrieved images for Occlusion affected based query images. Here, extra unintended and unwanted occlusion lines are drawn on both trained and testing images, by using DLCNN feature learning with normalization output images are retrieved accurately. And finally, Figure 5b represents the scenario of retrieved images for Noise affected based query images. Here, the proposed deep learning approach by default eliminates all types of noised and improves the retrieval accuracy

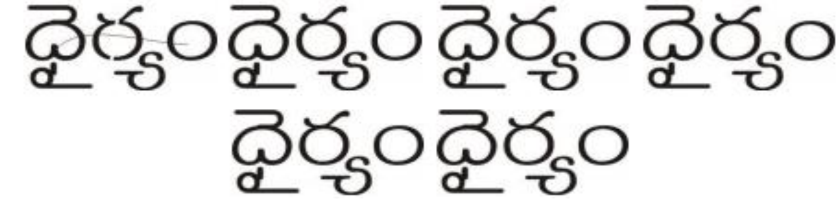
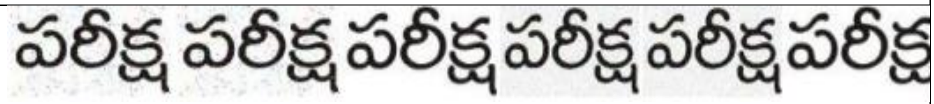
	Test Query word	Retrieved Telugu words
(a)		
(b)		
(c)		
(d)		
(e)		
(f)		
(g)		

Fig. 5: Retrieved output Telugu word images with DLCNN and DeepCluster (a) Missing segment with random distortion, (b) Noisy as a query word, (c) Occlusion effected, (d) Random disturbance and (e), (f), (g) Missing segment words.

Table 3: Performance comparison of proposed TWIR withexisting systems

Measurement	SIFT +	HMM-C	SURF+	GLCM-	HWNET	SDM-	Proposed

	BoVW [12]	[13]	BoVW [14]	IPC [15]	v2 [16]	NSCT [17]	TWIR system
mAP	0.853	0.89	0.731	0.967	0.978	0.998	0.999
mAR	0.799	0.823	0.809	0.842	0.92	0.98	0.999

Table 3 represents the performance comparison of proposed TWIR system using Deep learning-based feature extraction and feature clustering with the various literatures. From the table 3, it is observed that the proposed method gives the superior performance of mAP and mAR metrics compared to SDM-NSCT [17], HWNET v2 [16], GLCM-IPC [15], SURF-BoVW [14], HMM-C [13], and SIFT-BoVW [12] respectively.

5. Conclusion

This work presents the efficient mechanism of TWIR system using deep leaning based feature extraction and feature clustering approaches. Initially, the normalization process was applied, so all the test and training image sizes are perfectly adjusted to equal level. It helped to extraction justified texture features with reduced losses and distortions. Then, DLCNN architecture was applied to extract all types of word features based on the Telugu alphabet modeling. Then, AlexNet based DeepCluster method applied for effectively analyzing the words and classifying the features based on the Telugu grammar, thus words are indexed based on features with their grammatical attributes. Finally, pairwise hamming distance was utilized to find the similarity between query words to the database. The simulations are performed against various distortion scenarios such as missing segment with distortion, random distortion, noisy effected, missing segment in Query words. And in all those scenarios, the proposed TWIR method gives the better mAP and mARperformance evaluation compared to existing methods. This work can be extended to implement as the real time word search assistance system for Indian digital libraries by incorporating all the Indian languages.

References

1. Li, Ang, et al. "Generating holistic 3d scene abstractions for text-based image retrieval." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017. A
2. Unar, Salahuddin, et al. "Detected text-based image retrieval approach for textual images." IET Image Processing 13.3 (2019): 515-521. A
3. MK, Yanti Idaya Aspura, and Shahrul Azman Mohd Noah. "Semantic text-based image retrieval with multi-modality ontology and DBpedia." The Electronic Library (2017).
4. Zeng, Mengqi, et al. "CATIRI: An efficient method for content-and-text based image retrieval." Journal of Computer Science and Technology 34.2 (2019): 287-304.
5. Parcalabescu, Letitia, and Anette Frank. "Exploring Phrase Grounding Without Training: Contextualisation and Extension to Text-Based Image Retrieval." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. 2020.
6. Estrela, Vania Vieira, and Albany E. Herrmann. "Content-based image retrieval (CBIR) in remote clinical diagnosis and healthcare." Encyclopedia of E-Health and Telemedicine. IGI Global, 2016. 495-520.
7. Zaidi, Syed Ali Jafar, et al. "Implementation and comparison of text-based image retrieval schemes." International Journal of Advanced Computer Science and Applications 10.1 (2019): 611-618.
8. Unar, Salahuddin, et al. "A decisive content based image retrieval approach for feature fusion in visual and textual images." Knowledge-Based Systems 179 (2019): 8-20.
9. Zhou, Wengang, Houqiang Li, and Qi Tian. "Recent advance in content-based image retrieval: A literature survey." arXiv preprint arXiv:1706.06064 (2017).
10. Hossain, Md Shahadat, and Md Rafiqul Islam. "A new approach of content based image retrieval using color and texture features." Current Journal of Applied Science and Technology (2017): 1-16.
11. Latif, Afshan, et al. "Content-based image retrieval and feature extraction: a comprehensive review." Mathematical Problems in Engineering 2019 (2019).
12. R. Shekhar and C.V.Jawahaar, "WIR Using Bag of Visual Words", IAPS Int. Workshop on DAS, Gold Cost, QLD, Australia, pp. 297-301, 2012.
13. Nagasudhha.D, Y.M.Lattha, "Keyword Spotting using HMM in Printed Telugu Documents", In Proc. of Int. Conference on SCOPES, Paralakhemundi, India, pp:1997-2000, Oct. 2016.
14. N.Jayanthi, S.Indhu, "Inscription IR Using Bag-of-Visual Words", IOP Conf. Series: MSE, 225(1), pp.1-8, 2017.
15. K. M. Lakshmi and T. R. Babu, "A New Hybrid Algorithm for Telugu Word Retrieval and Recognition", Int. Jour. of IES, vol. 11, no. 4, pp.117-127, 2018.
16. P. Krishnaann, C.V.Jawahaar, "HWNET v2: An efficient word image representation for handwritten documents", Computer Vision and Pattern Recognition, 2018. arXiv:1802.06194 [cs.CV]
17. K. M. Lakshmi and T. R. Babu, "Robust algorithm for Telugu word image retrieval and recognition", Journal of Mechanics of Continua and Mathematical Sciences, vol. 14, no. 1, Feb. 2019.

18. Lakshmi, K. Mohana, and T. Ranga Babu. "A Novel Telugu Script Recognition and Retrieval Approach Based on Hash Coded Hamming." International Conference on Communications and Cyber Physical Engineering 2018. Springer, Singapore, 2018.
19. Pala, Mythilisharan, Laxminarayana Parayitam, and Venkataramana Appala. "Real-time transcription, keyword spotting, archival and retrieval for telugu TV news using ASR." International Journal of Speech Technology 22.2 (2019): 433-439.
20. Cheekati, Bindu Madhuri, and Roje Spandana Rajeti. "Telugu handwritten character recognition using deep residual learning." 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC). IEEE, 2020.