A novel automatic system for logo-based document image retrieval using hybrid SVDM-DLNN

Kudamala Raveendra*
Koneru Lakshamaiah Education Foundation (K L University),
Green Fields, Guntur District, Vaddeswaram,
Andhra Pradesh 522502, India
and
ECE Department,
Sri Venkateswara Engineering College for Women,
Tirupati, Chittoor, Andhra Pradesh, 517507, India
Email: raveendrasvew@gmail.com
*Corresponding author

R. Vinothkanna
ECE Department,
Koneru Lakshamaiah Education Foundation,
Guntur, Andhra Pradesh, 522502, India
Email: rvinothkannaphd@gmail.com

Abstract: Many government and private organisations represent themselves to the public using their own symbols or logo which is unique from others so that anyone can easily identify their products or belongings. This gives an ownership and source documentation to the owner by simply providing such logos. Using these logos for document retrieval in World Wide Web is a booming research in present era. Since usage of virtual documentation is increased day by day and handling this large data becomes a problem while searching for single data. In present research arena various document image retrieval models is available based on classification and clustering techniques. In this graph techniques are used to identify the issues in the automatic logo detection model using back propagation neural network along with the single value decomposition model. This proposed research model concerned about the document retrieval system based on the logo matching process to attain better efficiency and accuracy than the earlier detection models.

Keywords: logo recognition; detection; segmentation; document retrieval; feature extraction; logo extraction; feature matching; automatic system; computer aided; engineering; technology.


Biographical notes: Kudamala Raveendra has finished his Bachelor of Technology in Sagi Rama Krishnam Raju (SRKR) Engineering College, Andhra University, Bhimavaram and Master of Technology in Jawaharlal Nehru Technological University (JNTU) College of Engineering, Kakainada. He is working as Associate Professor in the Department of Electronics and
1 Introduction

The development of science and electronic media in the present life needs information access management system in every application in the format of e-documentation. This digital documentation models increases day by day since the usage of customers also increases rapidly. Many public and private sector libraries and many archives are interested in digitalisation of their documents and transcription of resources for better administrative and filing procedures. Since these documentations are paper based and when it is included in the ubiquitous digital environment their life will be increased as an unpredictable and non-destroyable one which provides better facilities than the paper format. The main objective of the documentation in the digital format is to process the various kinds of documents in a single place. This provides easy access to the users to collect all the information in single place not only for the single user and it is useful for large number of users. In traditional document retrieval systems an index will be used to obtain the document and later a high performance optical character recognition sheet is used to identify the particular document. This process works based on the ASCII and the provided data is compared using string comparison process to obtain the correct match of the actual document. But in this case OCR fails sometimes due to document degradation. This issue occurs in the OCR system due to degree of degradation which the document is accesses using scanner. Many document retrieval models are available in the current trend. In this logo based document retrieval is new and important in the research issues.

Logo is a simple graphical representation or an emblem of the organisation which may be public or private which is used for commercial purpose. This graphical symbols or graphics has advantage than any other advertisement and the public can easily capture the product names based on their logo. Generally logo includes name of the organisation or an indicating text which represents the firm. Some of the logo includes the organisation details and based on that logo is classified into
simple text logo

graphical logo

mixed logo.

In this simple text logo includes only the name for representing the company or their products. Considering Figure 1 Google, NASA is the examples for the simple text logo. Second graphical logo includes an image for representing the product or about the organisation information. Apple logo used in ‘i’ products and Pepsi logo in Figure 1 is an example for the graphical logo representation. In the last mixed logo is the combination of both text and graphical representation. KFC and burger king as seen on Figure 1 is the example for mixed logo. Apart from emblem logo is one kind which includes both text and graphical image along with the organisation name and details. Harvard logo in Figure 1 is an example for this emblem logo. Many organisation and enterprises used these kinds of logo to represent the ownership and authentication over the decades.

Generally logo analysis is divided into two steps such as detection of logo from the document image and classification of selected logo based on the data base. In this first step is used to detect the logo and second step is used to recognise the logo and that process is given logo recognition. Figure 1 provides a collection of various types of logo as an illustration.

Figure 1  Illustration of various logos available (see online version for colours)

Apart from logo some of the organisation used seals as their representation by providing an impression on paper using wax or some other ink mediums. This seals are widely used in many places by individual person to many firms for their identification. Seals are used to appear on top of the document as a imprints. This presence provides an authentication to the others for verification and identification purposes. As like logo seals are also classified textual and graphical and it bears information about the organisation some times. In many administrative process this seals plays a vital role for authentication and securing the document as sealed items.

This seals are used to set by manual impression of inkpad for representing the text or the graphical items. Apart from that it includes a variable field number; sometimes it includes index numbers, serial number, and date. Figure 2 illustrates different types of seals which is available in present time. In many applications content based image retrieval is used to employ in document image retrieval for handling the large scale of
data requests provided by the user. Based on the user request a search process is initiated for identifying the actual document which is stored already in database. Since document image retrieval is a developing research field which has on demand opportunities to develop a secure system.

Figure 2 Illustration of different types of seals, (a) text with date (b) text using membership numbers (c) closed contour surrounding as text along with year of starting

(a)    (b)    (c)

Complex noise, signature, logos are used to impose the algorithms used to retrieve the document from the data base. The simple steps involved in document image retrieval using feature component selection, providing suitable matching algorithms and allocating suitable indexing to the documents through images. In general text is classified into two different types such as documents and terms. In this common paper, books, messages are comes under documents category. Word, phrase which is used within the document is given as terms. The existing techniques used this vector space representation, Euclidean distances for representing the terms in the document image retrieval process.

Figure 3 Simple process diagrams for document image retrieval system

A simple document image retrieval model is depicted in Figure 3 in this a query image is given to the filter section for removal of unwanted noise and then it is provided as a input to the comparator which compares the actual query image using matching algorithm. This algorithm compares the actual query image in the document which is saved already in their data base. So that based on the colour identification, symbols, text and other features the actual documents are retrieved from the data base. The collected documents are used to index in the documentation process and then it is given as final image. This simple process handles the data for small applications and difficulty level may varied if the document data base has large number of data and query queue is also in large number. The main aim of this research is to identify the logo detection based document retrieval system for extracting the better features in the retrieval models. In general segments are
used to identify the logo in recognition model for identifying the practical issues and in recognising the logo, a simple model is essential to recognise them and that model is also discusses in the proposed research model. In this research limit image segmentation and logo detection are considered as beyond the scope of the research paper.

2 Related works

A vast survey has been made about the document image retrieval process and its evaluation over the decade. Dixit and Shirdhonkar (2015) discuss about the concept of feature based recognition a model which is used in their research model for identifying the simple blocks. The application is developed in the research based on the detection algorithm and segmentation algorithms. Similarly in Le et al. (2013) a simple logo based detection system is proposed for identifying the actual data from the data base. The researcher presents this documentation model for both the logo regions and non logo regions of documents. Based on the number of window and its size the density of the document is computed and this spatial density based model provides better logo recognition of the document from the database.

Sahbi et al. (2013) reports about the issues in the logo model used which are described as local curvatures, distance or centroid. These models used full logo image for extracting the actual document from the server and also it used a global descriptors for image detection as shape descriptors. This researcher uses logo pictures which are affected by noise and other transformational images. Based on the Pham et al. (2011) the real world images cannot be applied for transformations. Since the validity of the proposed research model has extensive experimental results which high lights the challenges and issues present in the data. Over 20% of baselines procedures are followed in the research for identifying and matching the data and documents. Nurdiyanto and Hermanto (2016) and Roy et al. (2011) describe about the research issues in the class of complicated objects such as signature. The earlier works of signature based models are used to select the actual document using simple recognition model. In this research a case study has been used to investigate the Persian based signature recognition model. Simple component analysis and other properties are used to recognise the signature regions. In the same time some other researchers are used a pixel based image recognition model for retrieving the actual document (Zhao et al., 2017).

Rusinol et al.(2011) describes about the information retrieval model used to obtain the document in digital libraries. This efficient model used coding techniques for document retrieval process. The limitation in the research is about to linguistics, since the recognition is based on the coding process and it does not support the linguistics used in the document. Alaei et al. (2013) uses principle component analysis for extracting the features in the identification process. Based on the extracted features matching process is performed in the research model by utilising weights of the image set. This is used to reconstruct the image if the image is not visible. Belhallouche et al. (2016) describes about the watermark based logo extraction system for extracting the distorted images. The proposed model used a combined discrete wavelet transform model along with the stochastic resonance images for identifying the watermarks in the materials. In the last this designed logo is hided on the watermarked image.
The findings from Sun and Chen (2011) describes about the dynamic stochastic resonance (DSR) based blind watermark extraction. The feature of this framework has blind watermark extraction and discrete cosine transform. The iterative process of research work employs various coefficients for analysing the watermarked images and the hidden information model. Qi et al. (2010) reports about the SIFT features in logo detection and recognition even in the sports data set. The important factor in the SIFT (Nejad and Faez, 2012) has its distinctive key points for image retrieval. A simple visual vocabulary is used to is avoided in order to identify the features. The M-estimation and other matched features are considered for obtaining the spatial relationships. More or less many original and duplicate features are identified under critical situation for performing the spatial clustering. All the models present in the literature has a generic proximity for identifying the layouts and duplicates.

Keyvanpour and Tavoli (2013) describe about the recognition of individual objects from a scene based on the peculiarities and problems present in it. It has contextual information about the actual image which includes the necessary information about the issues on hands. A spatial pyramid based approach is used in the image categorisation process and that is not suitable for better recognition model. Geometry of objects is considered in the joint distribution constellation models used in Sahbi et al. (2013). The approach behind this model is practical in most of the cases and a complex function is used to represent the number of parts as units. It is difficult to recognise such functions for data retrieval by a common model so that it provides better secure model than the existing one. In this a group of local image is chosen as spatial model for improving the accuracy. The features are matched and then it is refined using cluster models such as semi local spatial systems.

Medjahed (2015) describes about the relation between the features and its pairs used in the recognition process. The researcher introduces an object based detection model and localised model which combines the segmented data and local patch data. In the proposed model of Wang et al. (2018) a histogram-based context feature selection model is designed for spatial proximity. In this also SIFT descriptions are highlighted to identify the neighbourhood function as shapes of the context image. Psyllos et al. (2010) discusses about the feature descriptor based spatial model which handles a pair of features. A distribution of features and its neighbours are pointed in the proposed model for document retrieval process. Soysal and Alatan (2011) provide information about the study of recognising logos. Pyramidal tree (Chaudhuri et al., 2016) is considered in the discussed model for obtaining pyramid of image in low resolution manner. Logo is extracted (Zhang et al., 2017) in the in second process and then it is recognised in the third process. Based on the skew angle (Nejad and Faez, 2012), feature extracted from the actual image is blocked in the bounding box of the logo.

The last part of the finding provides information about the methods to increase the logo detection and occlusion rates (Meng et al., 2013). The spatial spectral saliency is used to avoid the background clutter in the proposed model which increases the efficiency of the system. The issue in the model is sensitivity since it is very sensitive to all occlusions. An SIFT based multiple resolution model is discussed in Wang and Chen (2009) and it has limitation in terms of image resolution. The local features and other spatial configurations are considered in the proposed model for obtaining better results. Research work of Xia et al. (2016) provide information about logo detection and segmentation based on the invariant shape detectors and its matching algorithms.
A non textual based document image retrieval model is used in Rodriguez-Vaamonde et al. (2015) to evaluate the important features of the proposed model. This framework retrieves application types, features, language and other costs related to the model. Another neural network-based model (Chen et al., 2018) is used in the research model for successive application like recognition and classification. Since neural network (Tavoli, 2012) plays an important role in the entire image classification model which uses multiple layers for non linear transformations. Convolution neural network (Tang et al., 2017) provides better image classification and its applications are not limited to image recognition and classification tasks. Based on the above survey (Zhang et al., 2018) works the proposed model is designed using neural network along with the single value decomposition model. This proposed model provides better computation efficiency (Nejad and Faez, 2012) in terms of fewer errors in detection.

3 Proposed work

The proposed model is designed as a automatic logo detection model for document image retrieval using hybrid model by combining the single value decomposition model along with Deep Learning neural network model (DLNN). It has an intelligent model for retrieval which manages the content and structural queries of the users. A tightly formulated system models are used to analyse the documents for the request in the query line and it needs a effective retrieval mechanism. Using various entities and document retrieval process the development of image retrieval architecture is provided for the proposed model which used both the control data and actual raw data. The control of document analysis model is defined for the logical structures such as textual, geometrical features. Figure 4 depicts the deep learning neural network model. In this input layer, hidden layer and out layer is present. The input and output layers are linked to each other by means of the hidden layer and the process of update begins with some random weights assigned to each node.

Figure 4 Illustration of deep learning neural network
Combination of these two features provides better representation for the extracted image. The documents and objects are stored in the data base which has sufficient provisions in case of handling complex large size data. This complex image has lots of image and other features include tabulated results which must be placed securely for protecting the contents of data. The core has the ability to control the applications based on the queries. Consider an example which has the query about the formal issue and the architecture searches for the actual data from its data base. After some time based on the results which is obtained a final retrieval result is provided to the users. This retrieval results includes image matching results, rank numbers and the maximum similarity numbers.

Figure 5  Block diagram proposed hybrid BPNN model for document image retrieval using logo (see online version for colours)

In this architecture based approach the representation of document is maintained as a critical one since the developed tools in this architecture has knowledge about the image. The query language is combined for the image and it is better in case of providing better
A novel automatic system for logo-based document image retrieval

Document retrieval process. Using image and text analysis tool the results are combined to produce an efficient and user friendly query tool to provide better solutions for even complex queries. It has an advantage such as image texture, image geometry, content features such as OCR text, image logical information such as structure, labelling, relations etc., in the proposed architecture model a hierarchical based classifier is used along with the classifiers like BPNN to train the different features and samples. This DLNN is combined with single value decomposition model for obtaining the error propagation algorithmic model for the proposed system. Figure 5 provides an illustration of proposed model architecture.

The proposed hybrid model uses input and output layers which are equal in number to obtain the vector elements and its classes. The hidden layers and its classification performance are compared with the hierarchical model kNN by trial and error method. The classifier segment the regions into text and picture based on the threshold based algorithms. The kNN is used in the proposed model to compare the results since it has a simple classification model. It provides better results for non linear boundaries and for each unlabeled patterns the k nearest neighbours obtains the distances. This classification method is considered for estimation of probabilities based on the frequency class of the neighbours.

Consider a query image \( Q \) with \( m \) number of key points and the data set is defined in the proposed model is given as

\[ Q = \{i_k, j_k, n_k\} \quad \text{with} \quad k = 1, 2, \ldots, m \tag{1} \]

where \( k \) represents the key points.

Consider another document image \( S \) with \( p \) number of key points and the data set is defined in the proposed model is given as

\[ S = \{i_k, j_k, h_k, n_k\} \quad \text{with} \quad k = 1, 2, \ldots, p \tag{2} \]

The matching between the each key point is related into

\[ T = \{i_q, j_q, h_q, n_q\} \tag{3} \]

For all the values of \( T \) the key point is computed. In the earlier stage the proposed model is used to consider the query feature in order to obtain the distance between the actual image and the query image. Considering the nearest neighbours the distances are given as

\[ d_1 = \min_{k=1,m} \{dist_e(Q_q, S_k)\} \]

\[ d_2 = \min_{k=1,p,q\neq q} \{dist_e(Q_q, S_k)\} \quad \text{with} \quad \arg \min_{k=1,m} \{dist_e(Q_q, S_k)\} \tag{5} \]

where \( dist_e(Q_q, S_q) \) is the Euclidean distance between the key points in the descriptor space. Given an unlabelled test pattern, suppose that \( \{k_1, k_2, \ldots, k_m\} \) denotes the numbers of nearest neighbours for the \( m \) classes. The estimate of the posterior probabilities is obtained as:

\[ P(\omega|\mathbf{x}) = \frac{k_i}{K} \tag{6} \]
In the classic KNN algorithm, x is classified in the class \( m \) if its posterior probability is the largest, i.e.,

\[
    c_m = \arg \max \omega_i P(o_i | x)
\]

In the case of LCO-ver1, the two most probable classes are found by choosing the two largest \( \omega_i \), which are equivalent to the two most frequent classes among the neighbour instances. For the second version of LCO, all \( \omega_i > 0, i \in \{1...c\} \) are chosen as probable classes. The proposed model detection algorithm for document image based on logo is given as below.

### Aim to detect the document based on logo

1. Initial the module
2. Begin
3. Input: query logo
4. Output: extracted document based on logo
5. Pre-process the query image and its related documents.
6. In the logo candidate state apply SVD to each possible logo
7. Decompose the matrices into orthogonal matrices and diagonal matrices
8. Calculate energy of each possible candidate using equation

\[
    e_i = \frac{1}{mn} \left( \sum_{j=0}^{n} \sum_{k=0}^{m} [u(i, j) + s(i, j) + p(i, j)] \right)
\]

9. Obtain the maximum number of combinations based on threshold values to obtain logo details
10. Apply multi-level filters for obtaining best features.
11. The distance between the features are identified using sample pictures
12. If (Detected the logo details)
    a. Retrieve the actual document as per the query request
13. Else
14. Repeat the process until it reaches
15. Detected logo = Candidate from the pool with Max. (Ei).
16. Extract detected logo from the document image.
17. End.

### 4 Result and discussion

The proposed model results are obtained over number of simulations to attain the performance. On each iteration the image in the simulated user searches for the data based along with the label. The simulated users click the relevant logo to train the target neural network for their inputs. The user feedback of images is given as 1 for relevant and 0 for all the other images in the data set. The range varies from 0 to 10 and it depends upon the dataset and accuracy of the search session. The system generally provides a training set and the search starts at random value by selecting any image in the data set.
A novel automatic system for logo-based document image retrieval

along with the target class. This target classes are used in the simulation environment to attain specific results in the search from the starting point. Finding the best pair based on the input logo and their ideal target document are highlighted as below. On iteration the database has a document which consists of logo and the user selected the logo then the document related to that logo will be retrieved as per the requirements of the user. In this model the experimentations are performed on Intel Core i5 processor CPU 3.00 × 4 GHz. The logo recognition is performed for the logos available in the data set and in this experimentation a graphical logo, text logo, mixed logo all the three combinations are chosen. In that the text logo has 30 numbers, text and graphics mixture logo has 10 numbers and the complete data set consists of 110 images. The data set is obtained from the automation research from Maryland University. Each logo in the data set is different in pixel size which is small in some pictures and large in some pictures.

Figure 6 Data set used in the proposed model

(a)   (b)   (c)

The data set provides 110 individual logos and some of them is intentionally degraded so that the logo recognition system performance can be measured along with the error rate. The recognition process includes three normalisation processes such as feature extraction, rotating the image at the centroid, resize the image into fixed width and height. The logo detection rate is used to obtain the performance of proposed model and it is given as

\[
\text{Logo detection rate} = \frac{\text{Number of logo detected accurately}}{\text{Number of documents in data base}}
\]  

(8)

The performance of the hybrid SVDM-DLNN is analysed based on the results for the inputs which is processed in MATLAB 14.1. The comparison of learning rate for train set and test set is given in Table 1

<table>
<thead>
<tr>
<th>Learning rate</th>
<th>Train set</th>
<th>Test set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Loss (%)</td>
<td>Accuracy (%)</td>
</tr>
<tr>
<td>0.1</td>
<td>10.35</td>
<td>87.21</td>
</tr>
<tr>
<td>0.01</td>
<td>8.58</td>
<td>92.31</td>
</tr>
<tr>
<td>0.001</td>
<td>7.12</td>
<td>94.65</td>
</tr>
</tbody>
</table>

From the result it was observed that the data set m and n produces different percentage of classification rates for the various functions. The process takes 98s of time to compute and produce the classification results for the network.

The proposed model has to obtain the best matched documents based on the logo so that the hybrid neural network shown in Figure 7 is used to train the data set used in the proposed model. The input image is given into the terminal and then it is test along with the pre trained values and the layers provides the better results and in the last output layer
K. Raveendra and R. Vinothkanna provides the details about the actual documents which is related to the logo. An accuracy metrics for different learning rates has been tabulated in Table 2.

Figure 7  Snapshot of NN train tool used in the proposed work (see online version for colours)

Table 2  Accuracy metrics for different learning rates

<table>
<thead>
<tr>
<th>Learning rate</th>
<th>Precision (%)</th>
<th>Recall (%)</th>
<th>F-score (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>84</td>
<td>75</td>
<td>73</td>
</tr>
<tr>
<td>0.01</td>
<td>86</td>
<td>78</td>
<td>76</td>
</tr>
<tr>
<td>0.001</td>
<td>82</td>
<td>74</td>
<td>72</td>
</tr>
</tbody>
</table>

The performance of the model depends on the learning rate and Table 2 depicts the accuracy metrics for the proposed model with different learning rates. The range of the learning rate varies and the precision, recall rate and f score also varies. Table 3 provides a comparison of parameters for the proposed model with existing kNN model.

Table 3  Comparison of logo detection rate

<table>
<thead>
<tr>
<th>S. no.</th>
<th>Logo type</th>
<th>Number of documents</th>
<th>Detection ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>kNN</td>
</tr>
<tr>
<td>1</td>
<td><img src="image1.png" alt="Logo1" /></td>
<td>60</td>
<td>81</td>
</tr>
<tr>
<td>2</td>
<td><img src="image2.png" alt="Logo2" /></td>
<td>80</td>
<td>75</td>
</tr>
<tr>
<td>3</td>
<td><img src="image3.png" alt="Logo3" /></td>
<td>90</td>
<td>76</td>
</tr>
</tbody>
</table>

The implementation of this hybrid model for document image retrieval is compared with the existing kNN based image retrieval model. From the results it is observed that the proposed model provides better performance than the existing model. Figure 8 provides as comparative analysis for three logo categories used and its detection rates for existing kNN and proposed SVMD-DLNN model.
It is observed that the findings indicate that the existing kNN provides less performance by selecting wrong documents if the logo is disturbed due to external noises. In case of proposed model a minimum threshold value is used as single value for minimum decomposition so that it is classified after that using deep learning neural network. This combination provides better results than other models which are available in the research.

Figure 8  Detection rate comparison for proposed model (see online version for colours)

Figure 9  Memory utilisation analysis (see online version for colours)
Figure 9 indicates the memory utility of proposed algorithm in terms of memory utility. The proposed work is found to consume less memory for the feature vector by an average of nearly 44%. The final plot is the computation analysis of the proposed work in terms of computational analysis of the proposed algorithm.

Figure 10 Memory utilisation analysis (see online version for colours)

References
A novel automatic system for logo-based document image retrieval


