Implementing the Data Mining Approaches to Classify the Images with Visual Words

Annaluri Sreenivasa Rao, Attili Venkata Ramana, S.Ramakrishna

***Abstract:** In the recent times, image classification got a significant importance and is going through wide range of challenges. In this work, a unique system is built to reveal the information of the image based on its category and relevant labels associated with it. To accomplish this task different data mining techniques are used for image classifications using Bag of Visual Words (BoVW) feature extraction algorithm. This algorithm is constructed with the help of grey level features and along with some of the colour features. Grey level features include speed up robust features (SURF), maximum stable external regions (MESR) and improved colour coherence vector (ICCV). Apart from that different techniques involved in data mining are being used includes neural networks (NN), decision trees (DT), Bayesian networks (BN), discriminant analysis (DA) and K-nearest neighbour (KNN). These techniques are used to evaluate the datasets included with Corel-1000 and COIL-100. BN and DA outperformed other methods as they are achieved a specificity of 99.9%, sensitivity of 99.5% and accuracy of 99.4% for Coral-1000. Whereas the same experimented results reported to be 100%, 98.5% and 98.9% respectively for COIL-100.*

***Keywords:** Data mining, Machine Learning, Image Classification, NN, KNN, BN, DA*

I. INTRODUCTION

Image classification is playing a vital role in the area of computer vision and in the area of pattern recognition researches. The role of image classification played important role includes, image retrieval, video retrieval, classification, analysing the images, surveillance, navigation, remote sensing, biometrics, etc. In simple words, it is a process of assigning different number of images for a predefined semantic category. The classification procedures are considered as two groups: a) supervised classification and b) unsupervised classification. In the first case, the set different classes will be identified from the existing images. However, in the later case the classes are not defined or identified. In the first case the classification is based on statistical characteristics of predefined information classes using the training methods. The most probable set of characteristics of the image will define the class to which the image belongs to. Whereas, in the second case, the training methods are not used for image classification. However, they examine the great numbers of different pixels of the images that are unidentified and then they are made to form groups. These groups of images are used as clusters based on different characteristics[13]. Sometimes the image classification considered as one of the complicated process due to some of the following factors: appearance with different light conditions, Changes in position, size, structure, noise levels and sometimes due to blur contents as well [12, 13, 15, 21]. In general, the classification of images must be very flexible for a wide range of visual classes with different range of classes. Many machine learning and data mining methods are in use to present these challenges according to Spangler et al. [24].

In this paper, an experiment is carried out using five (5) data mining techniques towards multi-class image classification and by using the following two steps procedures includes feature extraction and classifier implementation. In the case of future extraction the algorithm presented by Elnemr [6] was implemented, due to which the key points and mapping of the key points was possible. These were then converted into visual words. In the recent times, the key points and key regions of an image are used to retrieve and classify the images as they represent the local information of an image [4, 6, 15, 17, 29].These local points help to detect the highlighted points automatically using different types of detectors and are characterized by descriptors [14, 16].These descriptors are collected in large numbers to form a cluster of similar characteristics. Here each cluster is considered to be a visual word representing a local pattern which is formed by descriptors of a cluster.This in turn helps to identify the visual-word vocabulary which represents the local image patterns.Therefore, the images are characterized as vectors with a count of visual words, which in turn used as feature vector to classify the images. Therefore in the image classification, the role of vectors holding the count of different visual words plays a vital role. To build a BoVW model, speeded up robust features (SURF) and maximally stable external regions (MSER) descriptors plays a key role when they are combined together with color correlograms and improved color coherence vector (ICCV) featured vectors. The classification task will be done by providing these vectors as inputs to the NN, DT, BN, DA and KNN. The overall classification tasks are evaluated based on the huge data collected from the datasets which was collected earlier.

In this paper, a brief overview on the topics involved with machine learning and data mining techniques in Section 2. Some of the important concepts of the image classification is covered in Section 3. A procedure and the comparative study will be presented in Section 4 and the experimental results were shown in Section 5, followed by a summarized conclusion.

II. ROLE OF MACHINE LEARNING AND DATA MINING

The process of image classification needs a careful attention to have a best choice to finalize the end results,

Revised Manuscript Received on March 26, 2019

Dr. Annaluri Sreenivasa Rao, Department of Information Technology, Vallurupalli Nageswara Rao Vignana Jyothi Institute of Engineering &Technology, Hyderabad (Telangana), India.

Dr. Attili Venkata Ramana, Associate Professor, Department of ECM, Sreenidhi Institute of Science and Technology, Hyderabad, (Telangana), India.

Dr. S. Ramakrishna, Professor, Department of Computer Science, SVUCM & CS, SV University, Tirupati (A.P), India,

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

for which concepts of machine learning (ML) provide a best fit from the observed data and prior knowledge available at the hands [19]. Different algorithms from ML help to discover various complex patterns and descriptions from the example datasets. A complete new situation can be predicted from the outcomes of the ML techniques and these are originated from AI, statistics and different databases [28]. Once again the concepts of ML can be divided based on supervised learning and unsupervised learning along with other clustering techniques. Whereas in the case of data mining, it is a simple process which uses the data to discover the new patterns using different techniques of ML for making predictions and to explain the data. The data used in data mining will be obtained from different sets of data samples and / or earlier situations. Finally the output will be in the form of predictions towards a new set of examples or classes or categories related to the defined objectives.

The learning techniques in general consider the structural descriptions to explain the possible characteristics of the final output after processing the data. In most of the data mining applications, there are four types of data learning processes, which include: a) classification learning; b) association learning; c) clustering; and d) numeric prediction.

In the classification learning, a set of different classified examples from the concepts are expected to be learned for classifying the unseen examples. In the case of association learning different types of association among different features are considered to predict a specific class value. However, in clustering process the groups of similar values, parameters and types are formed to sort out the prediction process. Finally the numeric prediction is based on the numeric quantities and not based on a discrete class [28].

In the recent times, data mining delivered exemplary results as a tool to analyze and classify the images from the datasets, in large numbers and efficiently as well. In the next section different data mining techniques are discussed for an easy understanding.

III. III. TECHNIQUES INVOLVED IN DATA MINING

As discussed earlier there are different data mining techniques includes NN, DT, BN, DA and KNN classifiers.

A. Neural Network Classifiers

The terminology of neural network evolved from the human brain and based on the functioning of the neurons in a human brain. The neurons are connected in large numbers and they helps in establishing decisions based on the human observations (which is called as training in data mining technologies). The techniques involved in NN classifiers are data-driven and self-adaptive schemes which are capable of adjusting themselves based on the data with no additional support from the underlying model functions or specified descriptions. NN are capable of approximating different functions with greatest accuracy made them to be known for universal functional approximations.

The concepts of NN are explained by three basic functional layers as a) input layers; b) hidden layers; and c) output layers. Each layer is assumed to have different sets of groups of neurons and are interconnected with some of the weighted connections based on the input values and training values. In the hidden layers the weights are based on the training procedure and will randomly changes according to the sequence of datasets. The data to be processed is applied at the input layers and processed through the hidden layers, which functions based on the training. Finally the output data is retrieved from the output layers [1, 9, 30].

In the hidden layers, the individual examples go through a regular examination to train the network layers to generate the prediction results with a slight adjustment of the weights whenever it is needed to avoid the incorrect predictions. The process of adjustments within the layers takes place randomly and regularly until the desired criteria of prediction is met with the objectives defined. It takes times to get accurate predictions using NN since at the initial stages the weights of the neurons will be having some random values and they are supposed to be generating some nonsense resultant values at the output. Therefore, a kind of comparison between obtained values and the known values. The information obtained from this kind of comparisons helps to adjust the weights gradually. With the progress in training the network will be more accurate and generates the nearest values to the known outcomes. Sometimes, the networks even perform the role of an expert by predicting the unknown inputs as well once they are well trained for a long time with large numbers of input values. Therefore one can build wide range of predictive models using NN classifiers with reduced model structure and assumption (IBM SPSS, 2015). However, one needs to understand that the flexibility of NN does not mean to say that it is interpretable easily.

In this paper, multi-layer perceptron (MLP) concept is considered, which uses back propagation algorithm to train the connection weights of the hidden layers due to the robust nature and simplicity [9, 11, 28].

B. Decision Tree Classifiers

DT Classifier is considered to be one of the nonparametric data mining techniques represented with tree architecture (i.e. with different branches and leaves). In this architecture, a node is assumed to signify a test of an attribute, a branch is assumed to signify the resultant output from a test, and the leaves represent the predicted classes. Nodes are again classified into three types as root nodes, internal nodes, and leaf nodes.

In the DT classifiers a decision will be based on the hierarchical rule-based method which helps to select the path between root and internal nodes to reach a leaf. This total phenomenon represents a class of the image to be classified [9, 25].High dimensional training data can be manipulated easily by using the DT classifiers with no excessive support of designing and training requirements. Due to simple structure it allows easy interpretations and visualization of the prediction outputs. However, to overcome the problems with tree overfitting, two strategies are used to solve the problem. In the first case the expansion of the tree is halted for encountering certain criterion and in the second case the DT will be build by integrating the leaf nodes which helps in reducing the larger trees (i.e. tree pruning) [9]. In most of the classification problems, the decision trees are used with different models along with different splitting algorithms which enhance the purity of classes of the delivered data samples.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Some of the examples of decision tree models are given here: CHAID, ID3, Quinlan, C4.5, and C5.0. [7].Splitting the training samples in C5.0 decision tree model is based on fields providing maximum information gain. Therefore each subsample extraction is done by using C5.0 decision tree model in this work until the last subsample is not in a position to split again. Such splitting allows obtaining the lowest-level splits and after re-examination they are removed or pruned from the model as they do not contribute significantly to the models with proper values.

C. Bayesian Network Classifiers

This is considered to be a statistical classifier and is based on Bayes theorem, which constructs the probabilistic models by using posterior probabilities for predicting class labels of samples to be tested. BN classifier estimates the posterior probability using Bayes theorem from prior probabilities. It means that it calculates the probability of image features based on the training samples. This classifier also delivered high accuracy and speed for larger size databases as well [20].

BN consists of a network with various nodes representing an attribute for each node. These nodes are linked by direct edges with no cycles are also known to be directed acyclic graphs. Using the BN the known evidences such as common-sense, i.e. real time information can be combined to build a probability model. These models help to identify the likelihood of different occurrences based on the unlinked attributes. The graphical model of a BN will display different variables (i.e. nodes) from the dataset along with the probabilistic or conditional and independencies among them. The possibility of representing the nodes is possible in BN and the links (i.e. arcs) are not needed to be represented. Due to the robust nature of the networks the probability of missing information is very high and therefore predicting the information using the existing is very crucial.

The users will be able to understand the relationships between different features and clauses using BN helps to identify the problem areas and predict different effects of interferences. Apart from this the BN classifiers helps to provide efficient approach to avoid over fitting of the data and along with the clear visualization of the relationships involved in the networks.

Two methods with BN models based on Naïve Bayes model are mostly used in data mining techniques. They are Tree Augmented a) Naïve Bayes and b) Markov Blanket Estimation.

The first one is more effective to create simple BN models which are improved versions of Naïve Bayes models, since it allow each predictor to depend on adjacent predictors along with the target variables. The classification accuracy and favorable performance are the two major advantages of these models. However, due to simplicity many restrictions are being imposed on dependency structures which are uncovered among different nodes is the biggest disadvantage of these models.

In the case of Markov Blanket estimation, in the BN the set of different nodes consists of targets parents, children and other subsets as well. This method identifies different variables in the network which needs to predict target variables and it is capable of producing variety of complex networks with an extended time to produces the same.

D. Discriminant Analysis Classifier

DA is considered to be a multivariate statistical scheme which depends on the results obtained by the derivatives of linear equations to combine independent attributes or variables. These are discriminated effectively between different classes and the combination of these variables is considered to be discriminate function. This generate standardized and raw coefficients which are used to be the weight for identifying the best attributes for discriminating from different dependent groups [8, 23].

The DA classifier is best predictive model which suits for group membership and it is made up of different discriminant functions which are based on linear combinations of predictive variables. These predictive variables are helping to influence in providing better discrimination among the groups. The samples of various instances related with different group members are useful to produce the functions. Such functions are very useful to predict the new instances with similar measurements out of the predicting variables with unknown group memberships [11].

E. K-Nearest Neighbour Classifier

In this method the classification is based on the similarities as compared to other cases and is also considered called with instance-based learning. This concept KKN classifier development took place to recognize the patterns of data without the need of extractions from the existing stored patterns or cases. Similar cases and dissimilar cases are two types of cases depend on the distances of the cases. Therefore the dissimilarity of these cases is based on the distance. When a new class is observed means, the classification of that class will be based on the distances of the nearest model. The similar cases with nearest distances are considered as neighbors [11, 28].

This method differs with other image classification methods based on the time at which learning takes place. In this case the learning process depends on the measurement of the distance between adjacent classes and therefore it can be considered as a lazy process. However, in other methods the learning process takes place quickly and is eager to produce a decision and generate the resultant weights for neurons in the learning process.

The possibility of having one or more number of nearest neighbor is more in this method and the majority classes of closest KNN will be assigned to a new instance. Here the term K is representing the nearest neighboring classes with similarities. In this work the value of K is assumed to be 3-5 and Euclidian metric method is used for calculating the distance among the classes.

IV. THE PROPOSED APPROACH FOR COMPARATIVE STUDY

In this section, the proposed approach and procedure to carry out the comparative study based on different data mining techniques for image classification are considered in a systematic way.

A. Proposed Strategy For Comparative Study

In this work, feature extraction method was used to carry out the experiment for image classification and the same was applied with different classification approaches. BoVW model was implemented to obtain discriminative features of the images to understand the simplicity, efficiency and robustness. Local features are extracted using BoVW method and then for each training class the visual vocabulary classes were built based on the class type.

B. Implementation of Bo VW Model

This model depicts a document with some of the distinctive keywords set and this BOVW scheme is similar to bag-of-words technique. Because of the simplicity and by considering the image features as a set of words this method is extensively used by many for the purpose of image classification.

This algorithm proceeds in three steps: a) first step is focused to detect and extract the local features from the images; b) second step deals with construction of codebook (i.e. visual word vocabulary) using clustering methods based on the different features. This codebook is considered to be as a dictionary with detailed mappings of features and parameters involved with the set of images; and c) finally visual words frequency in the images are computed for image classification. The representation of the overall algorithm generates histograms with different visual features and occurrences which help in training the image category classifier to predict the image[6, 15].

Feature extraction tasks will be performed to build a dictionary with visual words and is carried out based on the Elnemr[6] proposals. The proposed algorithm detects SURF interesting points and MSER interesting regions as per the Elnemr's proposals. Later on the distinctive feature descriptors are being carried out to compute each and every point and region. The color correlograms and ICCV's are applied for extracting the color features, since SURF and MSER only operate with grey scale images.

SURF outperforms other detectors and descriptors due to very fast computational features and this algorithm is capable of representing the interesting points by a set of 64-dimensional descriptors. This was first proposed by Bay et al. [3] as a pioneering scale and rotation invariant key detector and descriptor. However, MSER detector extracts elliptical covariant regions using Watershed algorithm and these regions are known to be interesting regions. From each interesting region a set of descriptors will be extracted using SURF technique. So for all images there will be a set of interesting regions and set of key points represented by 64 dimensional descriptors. The advantages of MSER include variable size and can be computed globally for the total region. This technique is not limited with patch size, window size and the regions are invariant to affine transformations.

Similarly, the color features are extracted by applying the color correlograms and ICCV [5, 10]. Color correlograms generally refers to a color correlation of an image which is a function of corresponding spatial distances. The major advantage of this method is to capture pixels of color distribution in the form of color histograms with the spatial information of an image. Based on the quantized color used for feature extraction the size of the color correlation feature depends. At the time of exertion RGB color model will be applied and 64 quantized colors are implemented with two different types of distances. Hence a vector size of 2X64 is the feature of the correlograms.

In case of ICCV extraction, the featured vector is dived into two constituents: a) coherent and b) non-coherent Constituent. In the first case they hold the pixels which are spatially connected and in the later case it includes the pixels which are separated. This type of extraction used more spatial information as compared with conventional color coherence vector, so that it helps to improve the performance without the need of extra computations [5]. The ICCV feature vector found to be depicting 64 coherence pairs, included with both coherent and non-coherent pixels of a specific color from RGB space. Feature vector size of ICCV is also similar to the previous extraction method, i.e. 2X64.

The set of key point descriptor of each image vary in cardinality produces a kind of complex environment or difficulties for the learning techniques (classifiers) which in general need feature vectors with similar dimensions. Therefore a need of quantization of extracted local descriptors with respect to the feature space is needed into visual words (i.e. clusters) for forming a visual dictionary. To build a visual world vocabulary the K-means clustering algorithm is used to identify the local patterns of the images. The vocabulary size will be based on the number of clusters and as a result an image will be represented as a combination of vector of words from a document. For any image, the BoVW will be generated which holds the information about the occurrences of individual word and this information will be used as a feature vector at the time of image classification.

A. Procedure for Image Classification

For an image the BoVW model is constructed at the initial stages and then at the time of classification phase it is implemented for training and testing purposes. In this study different types of data mining techniques were considered for image classification to identify the best classifier which suits for giving a solution for the present problem. Also it is well known point to understand that one cannot consider any data mining method as a best one as compared to others as they are having independent pros and cons. Hence in this work several data mining techniques are tested and evaluate the results for finding the suitable image classification method.

B. Comparative Analysis

The examined datasets are used for evaluation strategy and different comparative studies are carried out using data mining algorithms. Comparative studies results help to portray the objectives for different data mining algorithms.

Dataset: Two different benchmark datasets are used for the comparative studies are considered from Columbia object image library (COIL-100) dataset and some of the images from COREL database [22 to 30].

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

With more than 100 different objects of color images are available in the COIL-100 database. About 72 views of the objects are obtained with different views of each object in different angles by rotating the objects. COREL-1000 data base is having a set of 10 irrelevant arbitrary real word classes, with a class of 100 color images from stock database of COREL photos. It contains different types of dinosaurs, cyber, houses, textures, dishes, elephants, fitness, Easter egg groups, bonsai, and other as well as shown in Fig. 1 and Fig. 2. The selection of these photos can be made randomly if needed and that provision is available.

Evaluation of the image classification methods is considered to be one of the lucrative topics in the area of experimental analysis which in general helps to permit objective selection and appropriate method to analyze the available data. The system performance is evaluated using the three following measurements: 1) accuracy, 2) sensitivity, and 3) specificity. The combination matrix is also considered for image classification and it is considered to be a cross-tabulation of different prediction values with respect to the observed values (i.e. true values) and accuracy depends on the empirical rate of correct predictions.

V. EVALUATION METHODS

Figure 1. Some of the Sample images from COIL-100 database

As discussed earlier the confusion matrix consists of a table which is applied to the classifier performance on a true set with the recognized data. Predicted classes of different input instances are shown by each column (row) of a matrix and the actual classes of inputs are given by each row (column). The following terms will be there in the confusion matrix, i.e. true positive (TP), true negative (TN), false positive (FP), and false negative (FN).

However, accuracy is known to be the empirical rate of correct prediction. The ability to classify the images correctly with respect to a particular class is known to be the sensitivity. Finally, the ability to predict images of other classes are not part of a predefined class is known to be specificity. These performance measures are computed using the following formulations:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Sensitivity = \frac{TP}{TP+FN} \quad (2)$$

$$Specificity = \frac{TN}{TN+FP} \quad (3)$$

A. Comparative Study Results

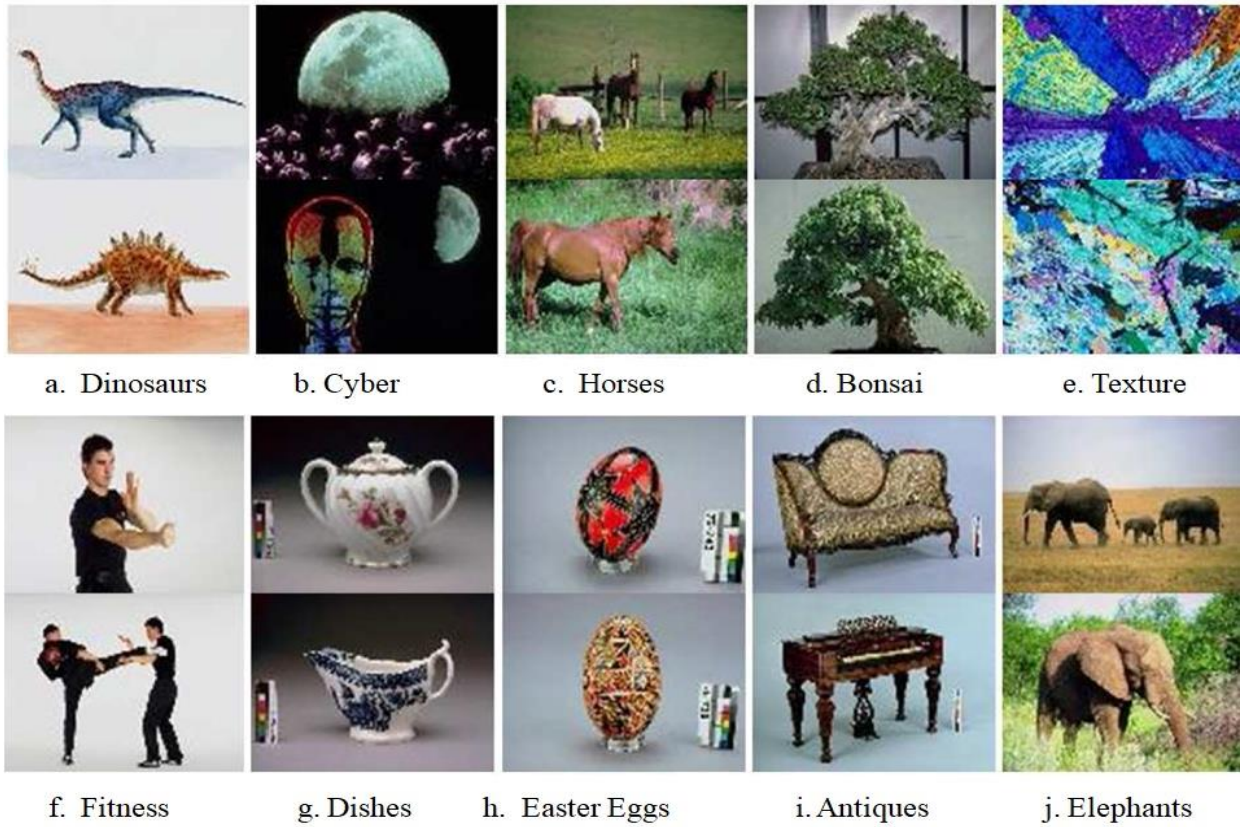
In this study, evaluation of the image classification using data mining techniques are studied and compared for NN, DT, BN, DA and KNN classifiers. All these classifiers are tested using BoVW model as proposed by Elnemr[6]. The obtained results for the images from COREL-1000 datasets are shown in Fig. 3, 4 and 5 for accuracy, sensitivity and specificity respectively.

Highest accuracy of 100% was achieved by using Bayesian classifier for dinosaurs, horses, bonsai and fitness classes of the figures as shown in Fig. 2. Also by using DA classifier for bonsai and antique classes achieved same 100% accuracy as shown in Fig. 3. Similarly, maximum sensitivity of 100 % was achieved by using Bayesian classifier for dinosaurs, cyber, horses, fitness, antiques, bonsai and elephants. Also by using DA classifiers bonsai, fitness, dishes, textures, dishes, Easter eggs, and antiques as shown in Fig. 4.The specificity also reported to reach to 100 % by applying Bayesian classifier for dinosaurs, horses, fitness, bonsai, dishes, and Easter eggs. However, also by using DA classifier for dinosaurs, horses, cyber, antiques, and elephants as shown in Fig. 5. The three measurements:

1) accuracy, 2) sensitivity, and 3) specificity are calculated as shown in Fig. 6, 7, and 8 describes the performance of overall system when applied to the COREL-1000 dataset.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

The results from the comparative studies shown that the BN and DA classifiers are found to be the best to use as image classifiers and outperformed other data mining techniques. An average accuracy and specificity of 99.9% are achieved by using both BN and DA classifiers. However, the sensitivity using BN and DA classifiers found to be around 99.5% and 99.4 % respectively. From this study it is proved that BN and DA classifiers performing excellently.



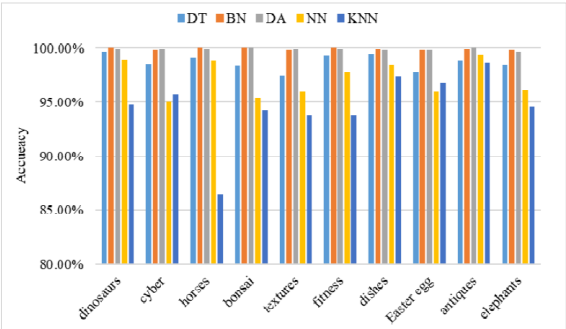


Figure 3.The data mining techniques applied on different classes of COREL-1000 dataset for accuracy values

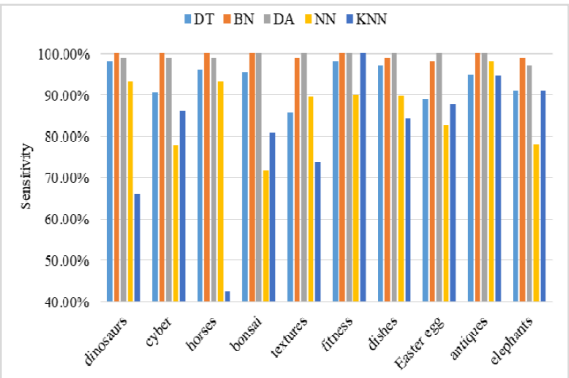


Figure 4.The data mining techniques applied on different classes of COREL-1000 dataset for sensitivityvalues

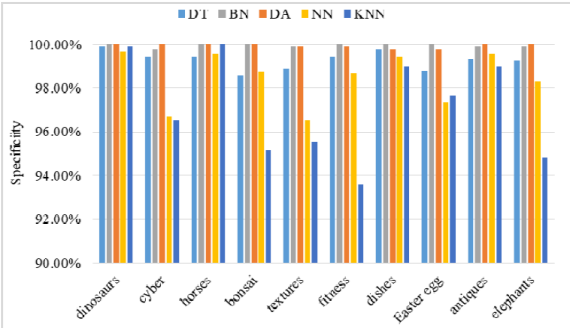


Figure 5.The data mining techniques applied on different classes of COREL-1000 dataset for specificity values

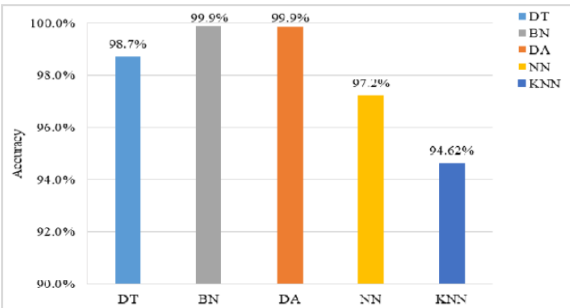


Figure 6.The average values after applying data mining techniques for different classes of COREL-1000 dataset

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

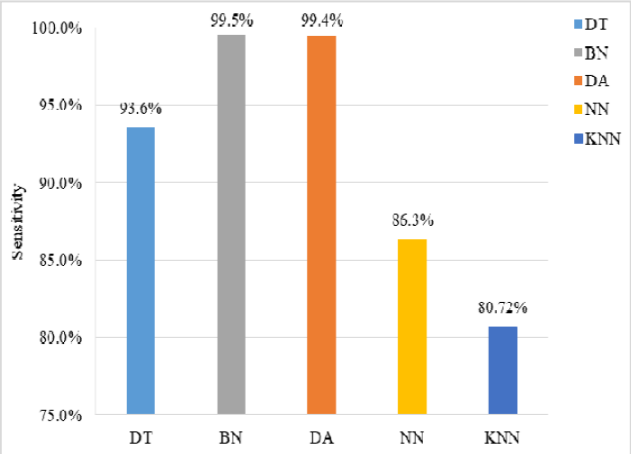


Figure 7.The sensitivity values after applying data mining techniques for different classes of COREL-1000 dataset

Average accuracy, sensitivity and specificity after applying data mining techniques for different classes of COIL-1000 dataset are shown in Fig. 9, 10, and 11. From the results it is seen that BN, DA and KNN classifiers are found to be delivering the best accuracy and sensitivity of 100%. In case of DT classifier it exhibited with an insignificant difference values of accuracy and sensitivity. However, the best values for specificity using DA classifier found to be 98.9% and by using the BN classifier the near optimal value found to be 98.5%.

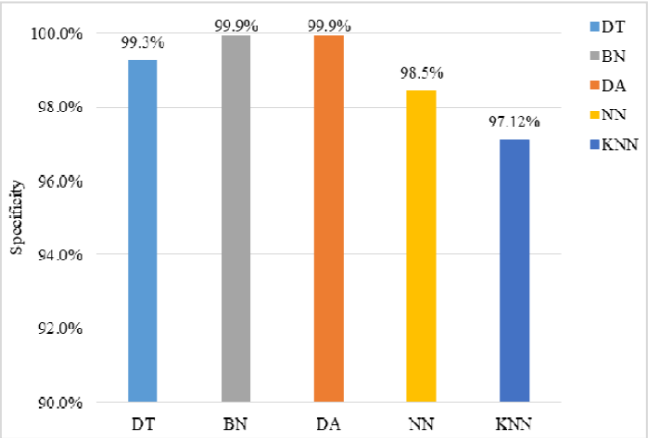


Figure 8.The specificity values after applying data mining techniques for different classes of COREL-1000 dataset

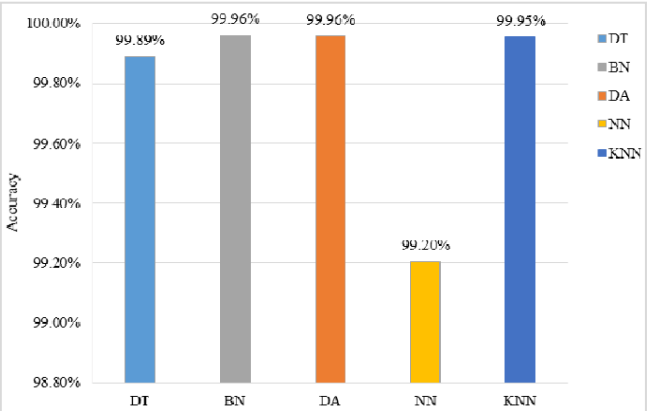


Figure 9.The average values after applying data mining techniques for different classes of COIL-1000 dataset

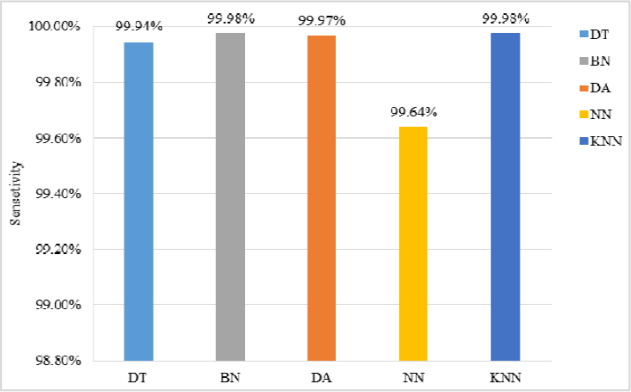


Figure 10.The sensitivity values after applying data mining techniques for different classes of COIL-1000 dataset

B. Discussion

The comparative studies carried out in this work revealed that data mining techniques are well suited for image classification and are well suited for handling the images intrinsically. DA and BN proved to be delivering best results for the image classification with better accuracy, sensitivity and specificity of near to 100%. On the other hand KNN proved to be delivering poor accuracy, sensitivity and specificity of 94.6%, 80.7% and 97.1% respectively.

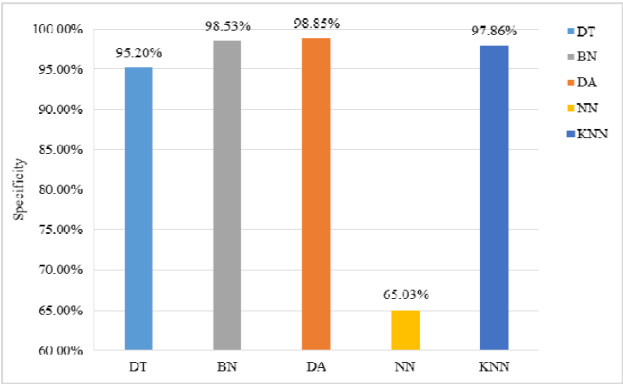


Figure 11.The specificity values after applying data mining techniques for different classes of COIL-1000 dataset

DA and BN proved to be independent models as a good prediction model and probabilistic models which allow them to give superior performances. It is seen that DA relies on discriminative function which help to combine different extracted attributes. Whereas, in the case of BN it depends on the degree of correlation of different attributes obtained. However, KNN does not support to build the classifier models and depends on the assigned value of neuron weights for all its attributes for computing the similarities between different images. This in turn lead to have more number of classification errors due to the changes in attributes on the regular basis and sometimes the probability of predicting a wrong image category is more.

Now the results obtained from the COIL-100 dataset for different data mining techniques revealed that DT, BN, DA and KNN delivered best accuracy and sensitivity of 100%,

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

whereas the specificity is 99% using BN and DA classifiers. Even in this dataset also the NN based image classification proved to be poor. However, for COIL-100 dataset DT and KNN proved to deliver the better performance as compared to the dataset of COREL-1000 due to uniform background in COIL-100, which is highly correlated with respect to the foreground objects.

VI. CONCLUSIONS

In this article, a detailed application of BoVW for image classification was tested for various data mining techniques. Different features (SURF, MSER, Color correlograms and ICCV) have been used to extract and build the needful dictionary of different visual words to train the hidden layers or to consider them as a reference for image classification from the unknown datasets as well. The featured vectors are identified from the images from different datasets and are fed to classification stage for retrieving image category.

Five data mining techniques DT, BN, DA, NN and KNN were investigated at the classification stage and compared in detail for accuracy, sensitivity and specificity. High level classification performances with nearest values of around 100% were achieved using the DA and BN data mining techniques. These two algorithms outperformed the remaining algorithms that were tested in this work. Images from different databases appeared to have their own impact on image classification due to the background and correlation factors. They impact on the performance of the classification and influence the overall results. For image classification the selection of the data mining technique to classify the images based on the type, nature and representation. Further the research may be extended to make an attempt to combine different data mining techniques to solve more complex images from different databases.

REFERENCES

1. Abe, B. T., O. O. Olugbara and T. Marwala, 2012. "Hyperspectral Image Classification Using Random Forests and Neural Networks". *Proceedings of the World Congress on Engineering and Computer Science*, October 24-26, San Francisco,USA.http://www.iaeng.org/publication/WCECS2012/WCECS2012_pp522-527.pdf
2. Acosta-Mendoza, N., A. Morales-González, A. Gago-Alonso, E. B. García-Reyes and J. E. Medina-Pagola. 2012. "Image Classification Using Frequent Approximate Subgraphs: In: Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications, Alvarez L., M. Mejail, L. Gomez and J. Jacobo (Ed.), Lecture Notes in Computer Science, vol. 7441, Springer, Berlin, Heidelberg, pp. 292-299. ISBN: 978-3-642-33275-3
3. Bay, H., A. Ess, T. Tuytelaars and L.V. Gool, 2008. SURF: Speeded up Robust Features. In: Computer Vision – ECCV 2006, Leonardis A., H. Bischof and A. Pinz (Ed.), Lecture Notes in Computer Science, vol. 3951, Springer, Berlin, Heidelberg, pp. 346-359. https://doi.org/10.1007/11744023_3
4. Benkrama, S., L. Zaoui and C. Charrier. 2013. Accurate Image Search using Local Descriptors into a Compact Image Representation. International Journal of Computer Science Issues (IJCSI), 10 (10): 220-225. https://www.ijcsi.org/papers/IJCSI-10-1-2-220-225.pdf.
5. Chen, X., X. Gu and H. Xu, 2007. An Improved Color Coherence Vector Method for CBIR. Proceedings of the Graduate Students Symposium of Communication and Information Technology Conference, (ITC' 07), Beijing.
6. Elnemr, H. A. 2016. Combining SURF and MSER along with Color Features for Image Retrieval System Based on Bag of Visual Words. Journal of Computer Sciences, 12 (4): 213.222. DOI: 10.3844/jcssp.2016.213.222
7. Fakhr, M. and A. Elsayad. 2012. Steel Plates Faults Diagnosis with Data Mining Models. Journal of Computer Science, 8 (4): 506-514. DOI: 10.3844/jcssp.2012.506.514.
8. Fernandez, G. C., 2002. Discriminant Analysis, a Powerful Classification Technique in Data Mining. In Proceedings of the SAS Users International Conference, April 14-17, Orlando, Florida, pp. 247-256. http://www.lexjansen.com/wuss/2001/WUSS01036.pdf.
9. Han, J. and M. Kamber, 2006. Data Mining: Concepts and Techniques. 2nd Ed. Elsevier, United States of America. ISBN: 13: 978-1-55860-901-3
10. Huang, J., S.R. Kumar, M. Mitra, W. Zhu and R. Zabih, 1997. Image Indexing Using Color Correlograms. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Jun 17-19, IEEE Xplore Press, San Juan, pp: 762-768. DOI: 10.1109/CVPR.1997.609412
11. IBM SPSS Modeler 17 Modeling Nodes. IBM Corporation, 2015. ftp://public.dhe.ibm.com/software/analytics/spss/documentation/modeler/17.0/en/ModelerModelingNodes.pdf.
12. Jain, M. and P. S. Tomar. 2013. Review of Image Classification Methods and Techniques. International Journal of Engineering Research & Technology (IJERT), 2 (8): 852-858. http://www.ijert.org/view-pdf/4730/review-of-image-classification-methods-and-techniques
13. Kamavisdar P., S. Saluja and S. Agrawal. 2013. A Survey on Image Classification Approaches and Techniques. International Journal of Advanced Research in Computer and Communication Engineering, 2 (1) 1005-1009. https://pdfs.semanticscholar.org/9509/c435260fce9dbceaf44b52791ac8fc5343bf.pdf.
14. Krig, S. 2014. Interest Point Detector and Feature Descriptor Survey. In: Computer Vision Metrics, Krig, S. (Ed.), Apress, Berkeley, CA, pp: 217-282. ISBN: 13: 978-1-4302-5930-5.
15. Liu, D., S. Wang and F. Porikli. 2016. Spatial Encoding of Visual Words for Image Classification. Journal of Electronic Imaging. 25 (3): 033008-(1-5). DOI: 10.1117/1.JEI.25.3.033008.
16. Lowe, D.G. 2004. Distinctive Image Features from Scale Invariant Keypoints. International Journal of Computer Vision. 60 (2): 91-110. DOI: 10.1023/B:VISI.0000029664.99615.94.
17. Lu, D. and Q. Weng. A survey of image classification methods and techniques for improving classification performance. 2007. International Journal of Remote Sensing, 28 (5): 823-870. Doi: 10.1080/01431160600746456.
18. Matas, J., O. Chum, M. Urban and T. Pajdla, 2002. Robust Wide Baseline Stereo from Maximally Stable Extremal Regions. Proceedings of British Machine Vision Conference, (mvc' 02), BMVA Press, Cardiff, Wales, pp: 384-393. DOI: 10.5244/C.16.36
19. Mitchell, T., 1997. Machine Learning. McGraw-Hill. ISBN: 0070428077
20. Murty, M. N. and V. S. Devi, 2011. Bayes Classifier. In: Pattern Recognition. Undergraduate Topics in Computer Science, vol. 0. Springer, London, pp. 86-102. ISBN: 978-0-85729-495-1
21. Naswale, P. P. and P. E. Ajmire. 2016. Image Classification Techniques- A Survey. International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), 5 (2): 236-239. http://www.ijettcs.org/Volume5Issue2/IJETTCS-2016-04-25-94.pdf
22. Nene, S., S. Nayar and H. Murase. 1996. Columbia Object Image Library (COIL-100). Technical Report CUCS-006-96, Department of Computer Science, Columbia University, New York. http://www1.cs.columbia.edu/CAVE/publications/pdfs/Nene_TR96_2.pdf
23. Ramayah, T., N. H. Ahmad, H. A. Halim, S. R. M. Zainal, and M.-C. Lo. 2010. Discriminant Analysis: An Illustrated Example. African Journal of Business Management, 4 (9): 1654-1667. http://www.academicjournals.org/AJBM.
24. Spangler, W. E., J. H. May and L. G. Vargas. 1999. Choosing Data-Mining Methods for Multiple Classification: Representational and Performance Measurement Implications for Decision Support. Journal of Management Information Systems, 16 (1): 37-62. http://dx.doi.org/10.1080/07421222.1999.11518233.
25. Tan, P.-N., M. Steinbach and V. Kumar, 2004. Introduction to Data Mining. Addison-Wesley Companion. ISBN-10: 0321321367

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

26. Wang, J., J. Li and G. Wiederhold. 2001. SIMPLIcity: Semantics-Sensitive Integrated Matching for Picture Libraries. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 23 (9): 947-963. DOI: 10.1109/34.955109
27. Vaishali, R., Sasikala, R., Ramasubbareddy, S., Remya, S., &Nalluri, S. (2017, October). Genetic algorithm based feature selection and MOE Fuzzy classification algorithm on Pima Indians Diabetes dataset. In *2017 International Conference on Computing Networking and Informatics (ICCNI)* (pp. 1-5). IEEE.
28. Reddy, S. R. S., Nalluri, S., Kunisetti, S., Ashok, S., &Venkatesh, B. (2019). Content-Based Movie Recommendation System Using Genre Correlation. In *Smart Intelligent Computing and Applications* (pp. 391-397). Springer, Singapore.
29. Rajasekaran, R., Kanumuri, U., Kumar, M. S., Ramasubbareddy, S., & Ashok, S. (2019). Sentiment Analysis of Restaurant Reviews. In *Smart Intelligent Computing and Applications* (pp. 383-390). Springer, Singapore.
30. Chugani, S., Govinda, K., &Ramasubbareddy, S. (2018, February). Data Analysis of Consumer Complaints in Banking Industry using Hybrid Clustering. In *2018 Second International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 74-78). IEEE.