ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 2, February 2014

Comparative Study of Supervised Classification Algorithms for WOSRAS

C. Chandra Mouli[1], P. Jyothi[2], K. Nagabhushan Raju[3]

Senior Research Fellow, Dept. of Instrumentation, Sri Krishnadevaraya University, Anantapur, AP, India [1]

Research Scholar, Dept. of Instrumentation, Sri Krishnadevaraya University, Anantapur, AP, India[2]

Professor, Dept. of Instrumentation, Sri Krishnadevaraya University, Anantapur, AP, India[3]

Abstract: A Wireless Object Sorting Robot Arm System (WOSRAS) is the combination of Machine Vision System (MVS), Wireless Embedded System (WES) and Robot Arm System (RAS). MVS is the essential fragment of the object sorting robot arm system which constitutes of an image sensor and LabVIEW installed personal computer system to classify the object from an image. NI vision acquisition express of LabVIEW acquires an image from the image sensor to identify the object in the image. The present work gives the intensive study of classification algorithms and the distance metrics. Each classification algorithm with its distance metric is evaluated to generate a classifier file. The performance of the classification algorithmsare compared to choose the best for object classification of image. NN (Nearest Neighbor), k-NN and Minimum Mean Distance (MMD) algorithms were considered for comparison. All the methods are analysed and compared for best results and maximum accuracy to implement in WOSRAS.

Keywords:Object Sorting, LabVIEW, RobotArm, Machine Vision System, Nearest Neighbor, Minimum Mean Distance.

I. INTRODUCTION

General procedure for sorting system constitutes three segments. They are Acquisition, Recognition, and Grasp and Sort. Acquisition is one of the important segments in sorting where reading of an image is done with a camera by focusing the image through effective lighting conditions. Identifying a specified object depending up on the nature of interest completes the recognition segment. Both the acquisition and recognition can be done using the software package installed in PC. Picking and placing the recognized object to a pre-defined place can be done by using a mechanical assembly completes the Grasp and Sort segment. This task is implementing in almost all places like libraries, pharmacies, warehouses, factories etc. In many sorting systems, the above steps are tied into one and cannot be separated clearly.

Sorting systems remain essential in numerous areas with diverse applications such as in manufacturing industry, libraries, factories, warehouses, pharmacies, supermarkets etc. MVS provides total automation solutions for industries to find defects, sort products and finish number of processes with speed and efficiently than humans. Yang Tao discusses the advantage of image processing in sorting applications by implementing a sorting system based on the hue extraction of an image processor from the image sensor and image processor performs a color transformation and obtains a single composite hue value for each object or piece of fruit to be sorted [1]. Thomas C. Pearson describes the object sorting system based on video image of an object [2]. Mohamad Bdiwi discusses about the control system and vision algorithms for library automation and book sorting using integrated vision/force robot control [3]. Roland Szabo implemented an object sorting system based on color using robot arm where web cam is used to identify the color of the object and robot arm is used to place the object in appropriate place [4]. A vision based robot system was developed for 3D pose estimation and picking of the objects in which a video camera surrounded by eight flashes is used to capture the images and CAD tool is used to find the edges of the object using a fully projective formulation [ACB98] of Lowe‘s model based pose estimation algorithm [5]. Raihan Ferdous Sajal and associates designed an efficient machine vision algorithm for real time image analysis and recognition of different features of Bangladeshi bank notes by using an automatic banknotes sorting system [7].

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 2, February 2014

II. LITERATURE REVIEW

Image Classification is defined as the process of extracting information/data from an image. The main role of image classification is to detect, recognize and classify the features of an object in an image depending on the type of class [8].The NN classification algorithmdetects the unknown object of a class in an image on the basis of nearest neighbor to the unknown classes form the trained classes. NN is the most widely used classification algorithm in ranking models [14], text categorization [11-13], pattern recognition [9, 10], event recognition [15] and object recognition [16] applications. kNN uses NN rulein which nearest neighbor is calculatedfrom the value of k to specifythe number of nearest neighbors to be considered to define a class of sample data point.kNN providesobjective, fast, transparent and produces good results over larger areas. The main advantage of kNN algorithm is its simplicity and lack of parametric assumptions [17].Past researches on Minimum distance classification showsthat it is extremelysuggested in all image classification applications because ofits minimum computation time as it mainly depends on the training data [21].

The present study was focused on the comparison and evaluation of classification algorithms to recognize the object of an image for WOSRAS. For this purpose NN, k-NN and MMD algorithms were studied, compared, and chosen the best classification algorithm.

III. SUPERVISED ALGORITHMS FOR IMAGE CLASSIFICATION

Classification is a decision-theoretic approach to identify the image or parts of the image. Image classification is one important branch of Artificial Intelligence (AI) and most commonly used scenario/method to classify the images among the set of predefined categories by using samples of a class.Image classification was categorized into two types; they are unsupervised and supervised image classification. Since the present work is based on the techniques of supervised classification algorithms rather than unsupervised classification algorithms.

Supervised classification is the most fundamental classification in machine vision classification. It requires prior knowledge of image classes. Training samples and test samples are used for classification purpose. An ordered pair (x, y) is called training samples where x is an instance and y is the label. An instance x with unknown label is called a test example. The aim of the supervised learning is to evaluate labels for test examples.

The most intuitive way of determining the class of a feature vector is to find its proximity to a class or features of a class using a distance function. Based on the definition of the proximity, there are several different algorithms, as follows.

A. Nearest Neighbor Algorithm

In NN classification algorithm,the input feature vector X of unknown class to a class $C_j$is determined as the distance to the nearestneighborwhich is used to represent the class as shown in (1).

$$d(X, C_j) = {}^{min}{}_i d(X, X_i^j) \quad (1)$$

Where, $d(X, X_i^j)$is the distance between X andX_i^j, the classification algorithm allots pattern X of unknown classto the class of its nearest neighbor that is given in (2)

$$X \in \text{Class } C_j, if\ d(X, C_j) = {}^{min}{}_i d(X, C_j) \quad (2)$$

The main advantage of the NN algorithm is its simplicity approach for classification. NN algorithm works well if corresponding feature vectors for every class are available.

B. k-Nearest Neighbor Algorithm

Depending on the voting mechanism the input feature vector X of unknown class to a class C_j is assigned in kNN algorithm. The classifier finds the k nearest samples from all of the classes. With the majority of the votes in the k nearest samples, the input feature vector X of the unknown class is allocated to the classC_j.

C. Minimum Mean Distance Classifier

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 2, February 2014

Consider that $\{X_{j1}, X_{j2}, \ldots, X_{jn_j}\}$ be n_j feature vectors which is used to represent the class C_j. Every feature vector contains a label of class j that has been selected for representing the class. The center of the class j is given in (3)

$$M_j = \frac{1}{n_j}\sum_{i=1}^{n_j} X_i^j \qquad (3)$$

An input feature vector X of unknown class was classified in theclassification phase depends on the distance to each class center and given in (4)

$$X \in \text{Class } C_j, if\ d(X, M_j) = {}^{min}{}_i d(X, M_i) \qquad (4)$$

Where,$d(X, M_j)$ is the distance function based on the distance metric selected during the training phase.

D. Distance Metrics

The performance of the classification methods were evaluated using LabVIEW. LabVIEW provides three distance metrics for the classification methods: Euclidean distance, Sum distance, and Maximum distance.Let $X = [x_1, x_2, \ldots x_n]$ and $Y = [y_1, y_2, \ldots y_n]$ be the feature vectors, then the distance metric $d(X, Y)$ was given for each distance metric.Euclidean distance (L2), Sum distance (L1) and Maximum distance (L∞) was evaluated and the resultant formulas for the distance metrics of classification methods was shown in TABLE I.

TABLE I
DISTANCE METRICS FOR CLASSIFICATION METHODS

Euclidean distance (L2)	$d(X, Y) = \sqrt{\sum_{i=1}^{n} (X_i - Y_i)^2}$
Sum distance, also known as the City-Block metric or Manhattan metric (L1)	$d(X, Y) = \sum_{i=1}^{n} \lvert X_i - Y_i \rvert$
Maximum distance (L∞)	$d(X, Y) = {}^{max}{}_i \lvert X_i - Y_i \rvert$

IV. METHODOLOGY

In the present work the classification algorithms along with its distance metric have been implemented and evaluated on two data sets i.e. standard data set and generated data set. Fig.1 shows the block diagram of the methodology for the evaluation of accuracy of classification algorithms. The training file, testing file, classification algorithm and distance metric is provided as input to the classification system. The training file contains the priori information regarding to the samples of each class in the quantitative form. Testing file is an image file contains image pixels which are to be classified. The classification procedure starts by assigning pixels to classes on the basis of similarity measures. Fig 2 shows the flowchartwhich describes the process of evaluation of classification techniques.By using confusion matrix and kappa coefficient the accuracywas evaluatedfor the classification algorithms.

Kohavi et al.in 1998 described about confusion matrix which consists of info about actual and prediction classes from the classification system [6]. Performance evaluation of the classification algorithms is done by using the information contained in the confusion matrix. The evaluation of classification algorithm's accuracy is the most common taskfor the assessmentof classification performance. According to Kohavi et al. Classification Accuracy, Misclassification Rate and Kappa Coefficient is given in (5), (6), and (7) respectively. The classification accuracy is defined as the proportion of total number of predictions that were correct.Misclassification rate (Error rate) is the opposite of classification accuracy.Another measure which can be extracted from a confusion matrix is the kappa(κ)coefficient [18] [19]. κ is defined as the extraction of correctly classified percentage from the actual

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 2, February 2014

percentage expected by chance.However the specified parameters for the performance evaluation are difficult to achieve in reality and no classification algorithm reaches the requirements nor applicable to all studies,because of data sets used and various environmental conditions.

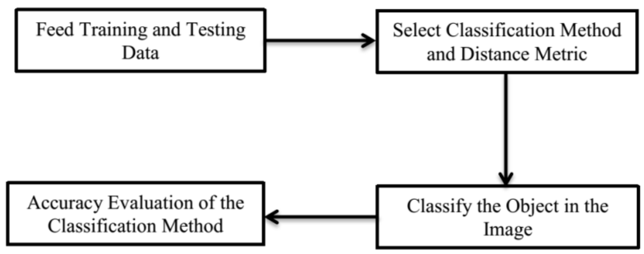

Fig. 1Block Diagram of the Methodology forAccuracy Evaluation of Classification Algorithms

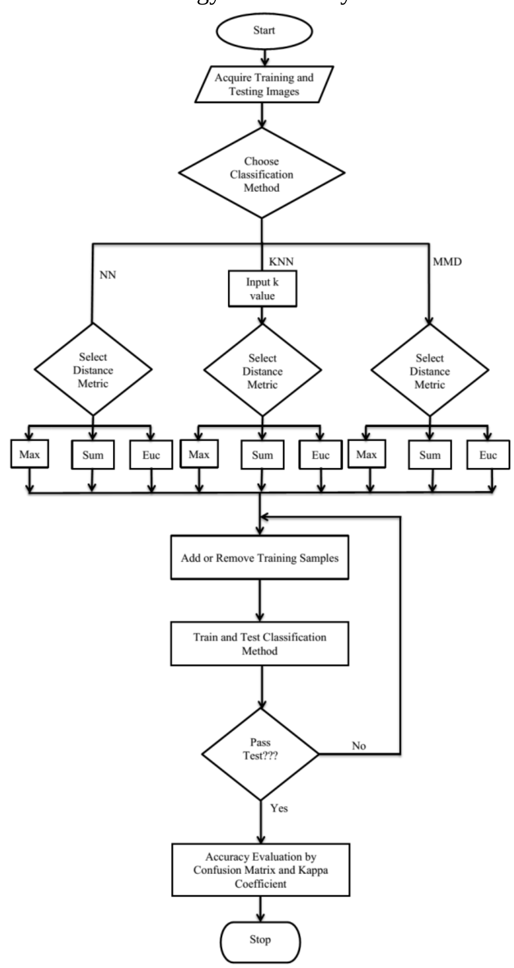

Fig. 2 Flow chart for the accuracy evaluation of supervised classification methods

$$Classification\ Accuracy\ (\%) = \frac{TP + TN}{TP + FP + TN + FN} \times 100 \quad (5)$$

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 2, February 2014

$$Missclassification\ Rate = \frac{FP + FN}{TP + FP + TN + FN} \quad (6)$$

$$kappa\ coefficient\ (\kappa) = \frac{P(D) - P(E)}{1 - P(E)} \quad (7)$$

Where, *TP* is the proportion of positive cases that were correctly identified, *FP* is the proportion of negatives cases that were incorrectly classified as positive, *TN* is defined as the proportion of negatives cases that were classified correctly,*P(D)* is given by the percentage of correctly classified instances (the sum of diagonal terms divided by the sum of total instances) and *P(E)* is given by the expected proportion by chance (the sum of the multiplication of the marginal probabilities per class divided by the sum of total instances).

NN/Max	Class 1	Class 2	Class 3	
Class 1	35	0	0	35
Class 2	0	33	2	33
Class 3	2	4	29	29
				97

(a)

NN/Sum	Class 1	Class 2	Class 3	
Class 1	35	0	0	35
Class 2	0	34	1	34
Class 3	2	1	32	32
				101

(b)

NN/Euc	Class 1	Class 2	Class 3	
Class 1	35	0	0	35
Class 2	0	34	1	34
Class 3	2	3	30	30
				99

(c)

kNN/3/Max	Class 1	Class 2	Class 3	
Class 1	35	0	0	35
Class 2	0	34	1	34
Class 3	1	4	30	30
				99

(d)

kNN/3/Sum	Class 1	Class 2	Class 3	
Class 1	35	0	0	35
Class 2	0	34	1	34
Class 3	2	3	30	30
				99

(e)

kNN/3/Euc	Class 1	Class 2	Class 3	
Class 1	35	0	0	35
Class 2	0	34	1	34
Class 3	2	3	30	30
				99

(f)

MMD/Max	Class 1	Class 2	Class 3	
Class 1	35	0	0	35
Class 2	14	4	17	4
Class 3	18	11	6	6
				45

(g)

MMD/Sum	Class 1	Class 2	Class 3	
Class 1	35	0	0	35
Class 2	12	14	9	14
Class 3	18	8	9	9
				58

(h)

MMD/Euc	Class 1	Class 2	Class 3	
Class 1	35	0	0	35
Class 2	12	10	13	10
Class 3	17	11	7	7
				52

(i)

Fig. 3 (a) NN with Max Distance Metric (b) NN with Sum Distance Metric (c) NN with Euclidean Distance (d) kNN with k value and Max Distance Metric (e) kNN with k value and Sum Distance Metric (f) kNN with k value and Euclidean Distance Metric (g) MMD with Max Distance Metric (h) MMD with Sum Distance Metric (i) MMD with Euclidean Distance Metric.

V. RESULTS AND DISCUSSIONS

The classification algorithms are tested and compared between the standard data set and the generated data set. The standard data set considered in the present work is IRIS plants data set. The accuracy and kappa coefficient for supervised classification algorithms for satellite images was evaluated and the results are given in [20]. The data contains three classes and it was randomly distributed into the ratio of 30:70 to obtain training and the testing file. The same criterion was used to implement the performance evaluation of the classification algorithms for WOSRAS.

The generated data set contains of three classes and it was randomly distributed into the ratio of 30:70 to obtain training and the testing file. Confusion Matrix for NN algorithm with Max Distance Metric, NN algorithm with

SumDistance Metric, NN algorithm with Euclidean Distance Metric, kNN algorithm with k=3 and Max Distance Metric, kNN algorithm with k=3 and Sum Distance Metric, kNN algorithm with k=3 and Euclidean Distance Metric, MMD algorithm with Max Distance Metric, MMD algorithm with Sum Distance Metric and MMD algorithm with Euclidean Distance Metric was shown in Fig. 3 (a), (b), (c), (d), (e), (f), (g), (h) and (i) respectively. Classification accuracy and kappa coefficient values for the classification algorithms are shown in Fig. 4.

	NN/ Max	NN/S Sum	NN/ Euc	kNN/3/ Max	kNN/3/ Sum	kNN/3/ Euc	MMD/ Max	MMD/ Sum	MMD/ Euc
Classification Accuracy (%)	92.38	96.19	94.2 8	94.28	94.28	94.28	42.85	55.23	49.52
Misclassification Rate	0.076	0.038	0.05 7	0.057	0.057	0.057	0.57	0.44	0.50
Kappa Coefficient (κ)	0.886	0.943	0.91 4	0.914	0.914	0.914	0.143	0.329	0.243

Fig. 4 Classification Accuracy and Kappa Coefficient Values

VI. CONCLUSION

Since the present study was focused on the comparison of classification algorithms to recognize the object of an image for WOSRAS. The different image Classification algorithms for object recognition namely Nearest Neighbor (NN), k Nearest Neighbor (kNN) and Minimum Mean Distance (MMD) along with the distance metrics Max, Sum and Euclidean have been implemented and tested with generated dataset. According to the results NN with Sum distance metric produces high accuracy very good kappa coefficient among the other classification algorithms. kNN with all distance metrics produces consistent and kappa coefficient. However, the accuracy of the algorithms depends on the robustness and quality constraints of training dataset. Different environmental conditions and selection of dataset also affects the classification accuracy.

REFERENCES

[1] Yang Tao, "Method and Apparatus for Sorting Objects by Color", U.S. Patent 5 339 963, Aug 23, 1994.
[2] Thomas C. Pearson, "Machine Vision Apparatus and Method for Sorting Object", U.S. Patent 5 703 784, Dec 30, 1997.
[3] Bdiwi, M.; Suchy, J., "Robot control system with integrated vision/force feedback for automated sorting system", Technologies for Practical Robot Applications (TePRA), IEEE International Conference , p p.31,36, 23- 24 April 2012.
[4] Szabo, R.; Lie, I., "Automated colored object sorting application for robotic arms", Electronics and Telecommunications (ISETC), 2012 10th International Symposium , pp.95,98, 15-16 Nov. 2012.
[5] Amit Agrawal, Yu Sun, John Barnwell, Ramesh Raskar, "Vision Guided Robot System for Picking Objects By Casting Shadows", International Journal of Robotics Research, vol. 29 no. 2-3, pp.155-173, February 2010.
[6] Raihan Ferdous Sajal, Mohammed Kamruzzaman, Faruq Ahmed Jewel, "A Machine Vision Based Automatic System for Real Time Recognition and Sorting of Bangladeshi Bank Notes", Proceedings of 11th International Conference on Computer and Information Technology (ICCIT 2008), pp.533-535, 25-27, December, 2008.
[7] B.Gabrya, L. Petrakieva; "Combining labeled and Unlabelled data in the design of pattern classification systems", International Journal of Approximate Reasoning , 2004.
[8] V.Vaidehi, S. Vasuhi, "Person Authentication using Face Recognition", Proceedings of the world congress on engg and computer science, 2008.
[9] Shizen, Y. Wu, "An Algorithm for Remote Sensing Image Classification based on Artificial Immune b-cell Network",Springer Berlin, Vol 40.
[10] G. Toker, O. Kirmemis, "Text Categorization using k Nearest Neighbor Classification", Survey Paper, Middle East Technical University.
[11] Y. Liao, V. R. Vemuri, "Using Text Categorization Technique for Intrusion detection", Survey Paper, University of California.
[12] E. M. Elnahrawy, "Log Based Chat Room Monitoring Using Text Categorization: A Comparative Study", University of Maryland.
[13] X. Geng et. al, "Query Dependent Ranking Using k Nearest Neighbor", SIGIR, 2008.
[14] Y. Yang and T. Ault, "Improving Text Categorization Methods for event tracking", Carnegie Mellon University.
[15] F. Bajramovie et. al "A Comparison of Nearest Neighbor Search Algorithms for Generic Object Recognition", ACIVS 2006, LNCS 4179, pp 1186-1197.
[16] B.K. Mayanka; "Classification of Remote Sensing data using KNN method", Journal Of Information, Knowledge And Research In Electronics And Communication Engineering(Volume – 02, Issue – 02).
[17] J. Cohen, "A coefficient of agreement for nominal scales", Educational and Psychological Measurement, 20:27-46, 1960.
[18] J. L. Fleiss. "Statistical Methods for Rates and Proportions",John Wiley & Sons, New York, 1981.
[19] Kanika Kalra, Anil Kumar Goswami, Rhythm Gupta, "A Comparative Study Of Supervised Image Classification Algorithms For Satellite Images", International Journal of Electrical, Electronics and Data Communication, ISSN: 2320-2084 Volume-1, Issue-10, Dec-2013.
[20] Kohavi, R., and Provost, F. "On Applied Research in Machine Learning", In Editorial for the Special Issue on Applications of Machine Learning and the Knowledge Discovery Process, Columbia University, New York, Volume 30, 1998.
[21] Richards, J.A., "Remote Sensing Digital Image Analysis: An Introduction", Springer-Verlag, pp. 265-290, 1995.