BENDLET TRANSFORM BASED OBJECT DETECTION SYSTEM USING PROXIMITY LEARNING APPROACH

Mritha Ramalingam
Faculty of Electrical and Electronics Engineering Technology,
University of Malaysia Pahang,
Pahang, Malaysia.
mritha@ump.edu.my

C.H. Nishanthi
Faculty of Science and Technology,
Bournemouth University,
Dorset, Bournemouth, UK.
chilukurinishanthi@gmail.com

Submitted: *May, 02, 2022* ***Revised:*** *Jul, 14, 2022* ***Accepted:*** *Jul, 30, 2022*

Abstract: This study presents a Bendlet Transform-based Object Detection (BTOD) system that recognizes an object in the image. Finding a specific object in images or videos is the goal of the field of object recognition. Though humans are able to identify a large number of objects, it is very difficult for computer vision systems in general. The appearance of the objects may change depending on the perspective, the size or scale, or translation and rotation. This work extracts Bendlet transform-based features from the images at different levels, and then the discriminant features are selected by employing genetic algorithms. The performance of the BTOD system is analyzed with different nearest neighbours for classifying objects in the Columbia Object Image Library (COIL-100) in terms of classification accuracy. It is observed from the results that the BTOD system with a one-nearest neighbour provides better performance than the two-nearest neighbour classifier. The former classifier gives 99.47% accuracy, whereas the later classifier gives 99.19%.

Keywords: Object detection, frequency domain transforms, spectral features, Bendlet transform.

I.INTRODUCTION

The number of images distributed and preserved is growing at an alarming rate all around the globe. The availability of such image content-related activity might complicate the transmission capabilities, the storage capacity of the hardware, and the amount of time required. When dealing with the indexing of a diverse and extensive collection of images, the automated categorization of images according to genre ought to be an effective strategy. A large amount of attention is paid to object detection compared to other types of image retrieval systems because of customers' needs.

In many computer vision systems, such as the navigation of robots, image retrieval, and object manipulation procedures, automatically recognizing objects in photographs plays an important role. The process of object identification by computers is a highly challenging one. Recent years have seen the development of many computer vision methods, which can identify or categorize objects in a scene. An efficient transform-based object detection approach is presented in this study to detect one hundred objects from various viewpoints. The remainder of this paper is organized into the following sections: The overviews of different object detection approaches are presented in Section 2. The development of a proposed system for object detection is explained in Section 3. The performances of the BTOD system are discussed in Section 4 statistically in terms of classification accuracy. The last section presents the findings and interpretations, which are based on the BTOD system's results.

II. RELATED WORKS

This section discusses several image processing approaches to detect objects from images. An object detection method using disparity values is discussed in [1]. The system utilizes the connected component algorithm and disparity values to connect pixels that have similar disparities. Additionally, the recursive object merge is used to merge similar neighbouring contours to locate and bound the contours of an object.

Supervised learning in a super-resolution network is accomplished using deep residual blocks for object detection [2]. It includes a Wasserstein Generative adversarial network. Then, the detection work is carried out using two cutting-edge detectors, Faster-Recurrent Convolution Neural Network and YOLOv3. An end-to-end blur-aid based object detection is performed in [3] using a feature aggregation network. It emphasizes the aggregation process impacted by the blur, including motion blur and defocuses. It also maintains a high accuracy level and requires just a little increase in computation.

A methodology is designed in [4], which categorizes automobiles in videos. It employs a histogram of gradient features and a linear Support Vector Machine (SVM) classifier as a machine learning technique. It also uses a YOLOv3 algorithm. Colour neural descriptors derived from convolutional neural networks by integrating several colour channels and colour spaces are discussed in [5]. These descriptors are based on the activations produced using a pretrained visual geometric group network.

An effective training method that gives a neural network with the dynamic-rate attribute is discussed in [6] for object detection. It consists of a neural network that can classify objects within the range of interest using a particular sensing matrix. The trained neural network only uses a small number of carefully chosen measurement rates for training. Many descriptors based on shape features are employed for object detection. An approach using Gaussian Hermite moments is implemented in [7] for object detection using the SVM.

The Empirical Wavelet Transform (EWT) is used in [8] to extract the characteristics of the objects. Energy and entropy are two characteristics that may be derived from EWT's parts. After that, the *k*-nearest neighbour classifier is used to identify the object. The wave-atom transform's energy features are extracted for object detection [9]. These features are used as input by the SVM classifier for the classification. Colour descriptors from different segments are taken into consideration in [10]. The use of edge mapping and clustering allows for the identification of distinct multi-coloured areas.

A colour vector representation model for object recognition is described in [11]. It uses Principal Component Analysis and Linear Discriminant Analysis to extract features from the data to generate the colour Eigenspace. The nearest neighbour algorithm has been chosen to serve as the classification mechanism. By using multi-resolution analysis and the discrete wavelet transform, separates the input into sub-bands to characterize the image for object detection [12]. It uses one nearest neighbor classifier for the classification.

III. PROPOSED SYSTEM

The proposed BTOD system uses the Bendlet transform [13] as feature descriptors and the proximity learning approach [14] as the predictor. It is an image processing-based system and is also considered a pattern (different objects) recognition system. The overall flow of the BTOD system is shown in Figure 1.

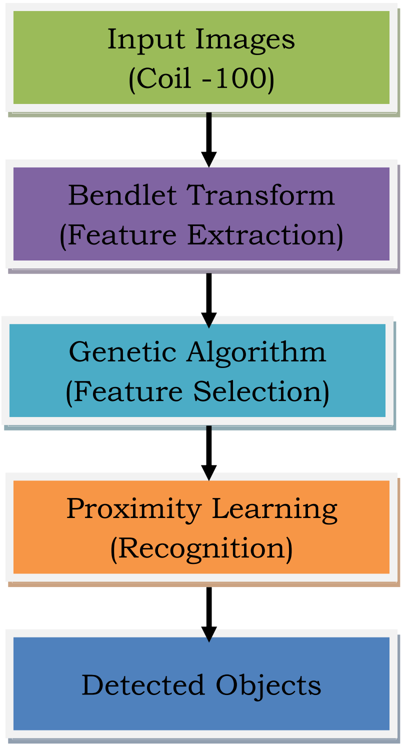

Fig. 1 Overall flow of the BTOD system

A. Bendlet Transform

It is highly desirable to pick a frequency domain analysis that can provide information ideal for recognizing objects in an image. The Bendlet transform is a Shearlet transform of the second order. The construction of a Bendlet is distinct from that of a traditional Shearlet in two aspects: the scaling operator and the shearing operator ($S_r^{(l)}$), respectively. In contrast to the alpha–scaling used by Bendlet, the traditional Shearlet uses parabolic scaling. In addition to that, the shearing operator is used with higher order. The Bendlet transform is defined for an input x as follows [12]:

$$S_r^{(l)}(x) := \begin{pmatrix} 1 & \sum_{m=1}^{l} r_m x_2^{m-1} \\ 0 & 1 \end{pmatrix} (x_1, x_2)^T \tag{1}$$

where l is the order of the shearing operator. The typical shearing matrix is obtained by setting l equal to 1. By setting l equal to 2, the Bendlet transform is constructed using a shearing element and a bending operator. The bending of the curves in the boundaries is classified into several categories by the bending operator. During the process of generating the scaling function, incorporating the anisotropy feature allows for the identification of the directional components that are present in the images. The alpha-scaling is

$$A_{a,\alpha} = \begin{pmatrix} a & 0 \\ 0 & a^{\alpha} \end{pmatrix} \text{ where } a > 0 \text{ and } \alpha \in [0,1] \tag{2}$$

To acquire the scaling matrix for Shearlet, alpha has to be set to 0.5, while Bendlet has to be set to below 0.5. Depending on the decay rate, the detection of borders might be very variable (Lessig et al. 2019). It has been shown that the application of the Bendlet transform results in an image with more precise curvatures.

B. Feature Optimization

Selecting features may be done more naturally with the help of genetic or evolutionary algorithms. One way to describe a feature subset is a binary bit string, often known as a chromosome. The value of each bit indicates whether or not the feature assigned to it is being utilized. Each chromosome stands in for a different location in the search space. Recombination, mutation, and selection are how the genetic algorithm brings about the evolution of a population of chromosomes in a manner analogous to the process of natural selection and the development of species [15]. In genetic algorithms, the recombination operator, also known as a crossover, is prioritized as the primary search operator, whereas mutation is used as a background operator with a low probability. They often use probabilistic selection (also known as proportional selection); most of the time, they depend on a binary representation. The fitness function used in this study is as follows:

$$f(x) = \alpha \cdot EG - mean + \beta\left(-\frac{|X|}{n}\right) \tag{3}$$

where the control parameters are α, β, and the number of features is represented by n and the number of features in the subset is $|X|$.

C. Recognition

The nearest neighbour classification approach used in this study for the classification process does not need any training [14]. The test data is best represented by the category based on its immediate *k*-nearest neighbours in one-nearest neighbor. The sequence in which the training samples are examined affects this outcome. In this investigation, the number of closest neighbours is set to 1 (k=1). A Euclidean distance measure is utilized to determine the distance between the training and the testing. The distance measure is defined as

$$ED(u,v) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \tag{4}$$

where $u = (x_1, y_1)$ and $v = (x_2, y_2)$ are training and testing features. Classifiers may either be soft or hard depending on the context. A hard classifier assigns a category to an item without providing an associated probability. It is assumed here that any object that satisfies the requirements for a certain class would, by definition, fall into that category. The classifier provides the probability of a soft classifier's categorization. The presumption here is that objects with seemingly identical characteristics may still belong to distinct classes.

III. RESULTS AND DISCUSSIONS

The performance of the proposed object detection system is tested on Columbia Object Image Library (COIL-100) database images [16]. Figure 2 shows the different objects in the COIL-100 database. A colour CCD camera equipped with a 25mm lens was fastened to a robust stand about 1 foot from the base of the stand. A motorized turn table was positioned around 2 feet from the stand's base. The turn-table went through its entire rotation of 360 degrees. Each object in this database has a resolution of 128x128 pixels and contains 72 images taken while rotating through five degrees. Figure 3 shows sample images in this database captured at different angles.

Fig. 2 Different objects in COIL-100database

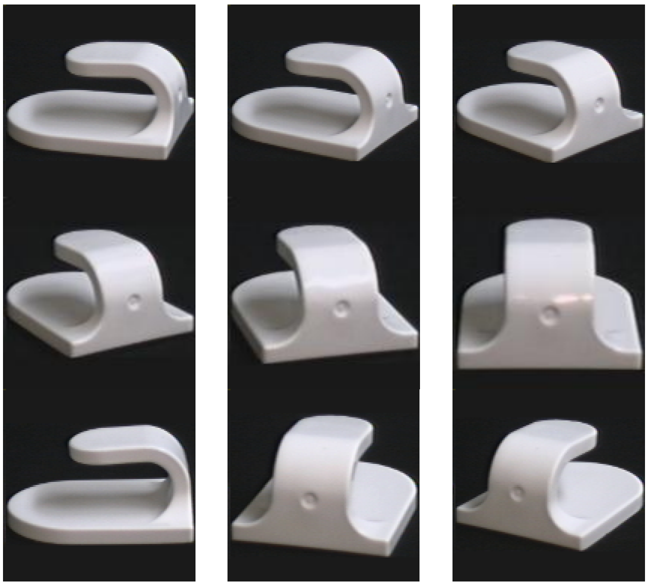

Fig. 3 Image captured with different angles of rotation

The performance of the proposed BTOD system is evaluated using classification accuracy. It is defined as the ratio between the number the objects correctly classified by the BTOD system and the total number of objects tested. A total of 7200 images are available in the COIL-100 database, and all the images are used for the evaluation. Table 1 and Table 2 clearly show the BTOD system's performance in classification accuracy for different levels of Bendlet transform.

TABLE. 1 Performance of the proposed BTOD system with 2-neighbour

level	Accuracy (%)				
	Ten	**Twenty**	**Thirty**	**Forty-five**	**Sixty**
1	93.17	89.00	85.05	79.08	73.44
2	95.33	91.83	88.25	82.11	76.33
3	96.92	93.70	90.25	85.03	80.93
4	98.53	95.91	92.75	88.22	81.99
5	99.19	98.20	96.50	92.92	89.46

TABLE. 2 Performance of the proposed BTOD system with 1-neighbour

level	Accuracy (%)				
	Ten	**Twenty**	**Thirty**	**Forty-five**	**Sixty**
1	93.17	89.00	85.05	79.08	73.44
2	95.33	91.83	88.25	82.11	76.33
3	96.92	93.70	90.25	85.03	80.93
4	98.53	95.91	92.75	88.22	81.99
5	99.19	98.20	96.50	92.92	89.46

The BTOD system's performance using a 1-neighbour classifier classifies the COIL-100 database image with an accuracy of 99.47%, whereas the 2-neighbour classifier provides 99.19%. It is also noted that for all combinations, the 1-nearest neighbour classifier offers more promising results than the 2-nearest neighbour classifier. Figure 4 clearly shows the results of 1-nearest neighbour graphically.

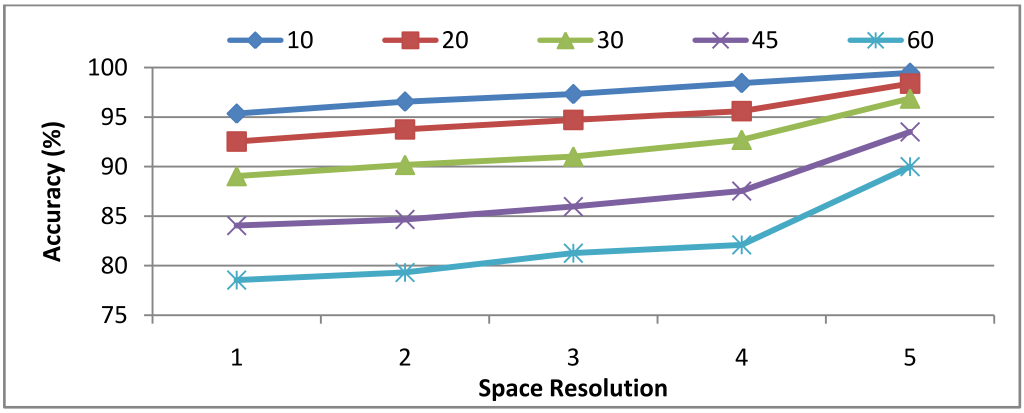

Fig. 4 Results of 1-nearest neighbour graphically

IV. CONCLUSIONS

This study presents and discusses an effective and efficient system for developing a computerized system for categorizing objects in an image. Bendlet transform is analyzed for feature extraction with a genetic algorithm, and the nearest neighbour technique is utilized for classification purposes. The genetic algorithm is employed for reducing the feature dimension of the Bendlet transformed image as Bendlet transform provides features in many sub-bands. It has been shown that the approach is quite successful for the automated categorization of objects. The system is evaluated using 7200 images from the COIL-100 database. The results obtained by the suggested method are highly encouraging, with maximum accuracy of 99.47%.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

REFERENCES

[1]. C. -C. Chiu and W. -C. Lo, “An Object Detection Algorithm with Disparity Values,” 4th International Conference on Imaging, Signal Processing and Communications, 2020, pp. 20-23.

[2]. L. Courtrai, M. -T. Pham, C. Friguet and S. Lefèvre, “Small Object Detection from Remote Sensing Images with the Help of Object-Focused Super-Resolution Using Wasserstein GANs,” IEEE International Geoscience and Remote Sensing Symposium, 2020, pp. 260-263.

[3]. Y. Wu, H. Zhang, Y. Li, Y. Yang and D. Yuan, “Video Object Detection Guided by Object Blur Evaluation,” IEEE Access, vol. 8, 2020, pp. 208554-208565.

[4]. A. S. Abdullahi Madey, A. Yahyaoui and J. Rasheed, “Object Detection in Video by Detecting Vehicles Using Machine Learning and Deep Learning Approaches,” International Conference on Forthcoming Networks and Sustainability in AIoT Era, 2021, pp. 62-65.

[5]. S. Saikia, L. Fernández-Robles, E. F. Fernández and E. Alegre, “Colour Neural Descriptors for Instance Retrieval Using CNN Features and Colour Models,” IEEE Access, vol. 9, 2021, pp. 23218-23234.

[6]. Y. Xu, W. Liu and K. F. Kelly, "Compressed Domain Image Classification Using a Dynamic-Rate Neural Network," IEEE Access, vol. 8, 2020. pp. 217711-217722,

[7]. B. C. Mohan, T. K. Chaitanya and T. Tirupal, "Fast and Accurate Content Based Image Classification and Retrieval using Gaussian Hermite Moments applied to COIL 20 and COIL 100," 10th International Conference on Computing, Communication and Networking Technologies, 2019, pp. 1-5.

[8]. S. Murugan, A. Bhardwaj and T.R. Ganeshbabu, “Fusion of Empirical Wavelet Features for Object Recognition,” International Journal of Engineering Research & Science, vol. 2, no. 12, 2016, pp. 99-105.

[9]. T. Gladstan and E.A. Mohan, “Novel Approach Object Recognition Using Efficient Support Vector Machine Classifier,” International Journal of Electronics and Communication Engineering and Technology, vol. 8, no. 2, 2017, pp. 81-90.

[10]. S.K. Naik and C.A. Murthy, “Distinct multicolored region descriptors for object recognition,” IEEE transactions on pattern analysis and machine intelligence, vol. 29, no. 7, 2007, pp.1291-1296.

[11]. A. A. M. Hagar, M. A. M. Alshewimy and M. T. F. Saidahmed, "A new object recognition framework based on PCA, LDA, and K-NN," 11th International Conference on Computer Engineering & Systems, 2016, pp. 141-146.

[12]. S. Elakkiya and S. Audithan, "Feature based object recognition using discrete wavelet transform," Second International Conference on Current Trends In Engineering and Technology, 2014, pp. 393-396.

[13]. C. Lessig, P. Petersen and M. Schäfer, "Bendlets: A second-order shearlet transform with bent elements", Applied and Computational Harmonic Analysis, vol. 46, no. 2, 2019, pp. 384-399.

[14]. K. Beyer, J. Goldstein, R. Ramakrishnan and U. Shaft, "When is "nearest neighbor" meaningful?" In International conference on database theory, Springer, Berlin, Heidelberg ,1999, pp. 217-235.

[15]. L.M. Du, Y. Xu and H. Zhu, "Feature selection for multi-class imbalanced data sets based on genetic algorithm," Annals of Data Science, vol. 2, no. 3,2015, pp. 293-300.

[16]. S.A. Nene, S.K. Nayar and H. Murase, "Columbia Object Image Library (COIL-100)", Technical Report CUCS-006-96, 1996, vol. 7, pp. 1-6.