A Comparative Study of Artificial Intelligence Algorithms Used for an Accurate Image Geometric Dimensions Recognition

Mahmoud Mohamed Hamoud Kaid, **Muawia Mohamed Ahmed Mahmoud**

AL Neelain University, Faculty of Engineering, Khartoum, Sudan
Assoc. Prof. Tabuk University, Faculty of Engineering, KSA.
(E-mail: Mahmoudkaid80@gmail.com, Muawia5858@hotmail.com)

_____

Abstract: The process of distinguishing objects in digital images and recognizing them is a basic process in AI, in order to distinguish the object and determine its features. It has many uses in various engineering and medical fields. There is a great difficulty in recognizing digital images and distinguishing objects in them. This paper discusses some basic artificial intelligence techniques based on the geometric dimensions of the object in the image in order to recognize these objects and distinguish them from the background of the image. Where there is great difficulty in separating the object from its background, this paper presents the most accurate methods of programmatic application using Matlab environment to identify components in the image, its distinction, and its characteristics. The paper provides a comparison between these AI methods and chooses the most accurate one. AI recognition methods use various algorithms including algorithms based on color density and algorithms that depend on geometric dimensions such as Hue moment algorithms, Haralik features algorithms, and Zernike moment algorithms. In this paper these algorithms were applied to a group of images to extract the features of the object and the best one will be chosen with the aid of Matlab. The most accurate recognition process is chosen through building a digital library that contains many pictures and training the program on these images to recognize and distinguish the object in the image.

Keywords: Artificial Intelligence (AI), Algorithms, Object Recognition, Hue moment, Zernike moment, Haralick Features.

_____

1. INTRODUCTION

The process of recognizing objects in the image is a significant area of AI. Detecting objects is one of the basic processes in AI as it provides important information about the objects in the image, such as where these objects are placed, their sizes, colors and other useful information. Therefore, this process is necessary for the operations of digital image processing that are built on AI algorithms in identification and discrimination processes. [1].

Some AI image detecting algorithms are based on features relate to color while there are other algorithms regard the geometry of shape such as Hu Moments, Haralick features, and Zernike Moments. This article discusses the AI techniques based on the geometric dimensions of the object in the image and also find out and presents the most accurate methods using Matlab program.

2. AI algorithms for Image Features Extraction

There are many algorithms AI that are interested in this area, including algorithms that relate to feature-color object within the image of which extract patches algorithm and algorithm Histogram Color which is a function where the division of the image to (patches) and extracts of which features that related by color and the distribution of these colors in the picture [1, 2].

The algorithms related by geometrical shapes of objects and the image are converted into a gray image formula which carry the Features of objects are extracts from the image to form information for the process of recognizing the object. This is done by comparison with the object images that were previously stored in the computer library.

The algorithm is then trained with the new Features that were the input to the algorithm [3]. Since these features are related With the geometry of the objects in the image, the algorithm will be trained to recognize, provide the decision, and classify these objects.

2.1 Hue Moments Algorithm

This algorithm is based on the calculation of (Image moments) which express the pixels belonging to the element inside the image according to a statistical equation that can form the number necessary to be used in Hu Moments algorithm where a specific number is formed and then combined according to a seven mathematical statistical equations to form seven features not affected by rotation, size or distance. Hence one can distinguish the objects [4].

2.1.1 Object Recognition by Geometrical Moment Invariants.

Hue defined seven descriptors that are computed from central moments in order of the three independents of the object: translation, scale and orientation. Translation invariance is obtained by calculating moments that are normalized and regarding the center of gravity [5, 6].

The size invariance can be computed from algebraic invariants. From the second and third order values of the normalized central moments, one can compute a set of seven invariant moments independent of rotation [5]. Given a two-dimensional image, $f(x, y)$, the moments of order (p, q) are defined in equation (1).

$$m_{pq} = \sum_{Z=0}^{M-1} \sum_{Y=0}^{N-1} z^p \cdot y^q \cdot f(y,z) \qquad (1)$$

$$p,q = 0,1,2,\ldots,\infty$$

Where: M, N are the horizontal and vertical dimensions of the image [8].

$f(y,z)$: Intensity (gray class) at the point (y,z) of the picture.

The zero-torque put p, q = 0 in the relation and giving the following equation:

$$m_{00} = \sum_{Z=0}^{M-1} \sum_{Y=0}^{N-1} f(y,z) \qquad (2)$$

The relation represents the total values of the pixels in the image and is named an area of the image. For binary image $f(y,z)$ the image area is equal to the flat area of the sample. It then determines the torque of the first class described in equations (3) [7, 8].

$$m_{10} = \sum_{Z=0}^{M-1} \sum_{Y=0}^{N-1} x\, f(y,z)$$

$$(3)$$

$$m_{01} = \sum_{Z=0}^{M-1} \sum_{Y=0}^{N-1} y\, f(y,z)$$

Sample Center is one of the important parameters used to determine the position of the sample. This is the point at which owns coordinates x', y' which is the sum of squares of the distance of each of the other points within the small object. The center of moments is expressed as:

$$y' = \frac{m_{10}}{m_{00}}, \qquad z' = \frac{m_{01}}{m_{00}} \qquad (4)$$

2.1.2 Central Moments for Features Object.

Corresponding to moments of inertia, moments samples depend on the location of the center and scale. The m_{20}, m_{02} studied sample are therefore of limited use. A collection of moments should be used to trace the changing. This group can be derived by calculating the central moments μ_{pq} given by the following equation:

$$\mu_{pq} = \sum_z \sum_y (z-z')^p\, (y-y')^q\, f(y,z) \qquad (5)$$

According to the previous equations (3) and (4), the moments are expressed as:

$$\mu_{10}, \mu_{01}, \mu_{11}, \mu_{10}, \mu_{20}, \mu_{02}, \mu_{12}, \mu_{21}, \mu_{03}$$

Hence:

$$\mu_{00} = m_{00} \quad , \mu_{10} = 0 \ , \ \mu_{01} = 0$$

$$\mu_{20} = m_{20} - z'.m_{10}$$

$$\mu_{02} = m_{02} - y'.m_{01}$$

$$\mu_{11} = m_{11} - y'.m_{10}$$

$$\mu_{12} = m_{12} - 2\, y'.m_{11} - z'.m_{02} + 2\, y'^2 .m_{10} \qquad (6)$$

$$\mu_{21} = m_{21} - 2\, z'.m_{11} - y'.m_{02} + 2\, z'^2 .m_{01}$$

$$\mu_{03} = m_{03} - 3\, y'.m_{02} + 2\, y'^2 .m_{01}$$

$$\mu_{30} = m_{30} - 3\, z'.m_{20} + 2\, z'^2 .m_{10}$$

Then the development of the central moments measured η_{pq} can be extracted as:

$$\eta_{pq} = \frac{\mu_{pq}}{(\mu_{00})^\lambda} \ ; \ \lambda = \frac{p+q}{2} + 1 \ ; \ p+q \geq 2 \qquad (7)$$

2.1.3 Object Features (Moments Hue Non-changing)

These parameters can define a set of non-changing moments as shown in equations (8) [9]. These moments non-changing form a complete set to describe the image and provide the possibility to recognize the samples with high accuracy [10].

$$\varphi\_1 = \eta\_20 + \eta\_02$$
$$\varphi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$
$$\varphi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$
$$\varphi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$
$$\varphi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) \begin{bmatrix} (\eta_{30} + \eta_{12})^2 \\ -3(\eta_{21} + \eta_{03})^2 \end{bmatrix}$$
$$+ \qquad (8)$$
$$(3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$
$$\varphi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] +$$
$$4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$$
$$\varphi_7 = (3\eta_{21} - \eta_{30})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})]$$
$$+(3\eta_{12} - \eta_{30})(\eta_{03} + \eta_{21})[3(\eta_{30} + \eta_{12}) - (\eta_{21} + \eta_{03})]$$

The torque account for a set of images of samples trained by the program, and saved in advance to be a library. The program compares the moments calculated for the current image coming from the image capture with moments stored, and based on the result of the comparison program recognizes or failure to identify the sample which is done by calculating the distance

2.2 Haralic Features AI algorithm

The Gray-Level Co-occurrence Matrix (GLCM) appears to be a well-known statistical technique for feature extraction. The GLCM is a formulation of how often different combinations of pixel gray levels could happen in an image.

The goal is to give an unknown sample image to one of a set of known texture classes. Textural features can be scalar numbers, isolated histograms, or empirical distributions. They characterize the textural properties of the images, such as a spatial structure, contrast, seediness, orientation, and have the certain correlation with the wanted output [11].

Gray level co-occurrence matrices (GLCM) suggested by Haralick have come to be one of the greatest well-known and generally used texture measures.

Let a two-dimensional image $I(x,y)$, $(x = 1, M, y = 1, N)$ have N_g gray levels. A co-occurrence matrix depicts the joint gray-level histogram of the image (or a region of the image) in the form of a matrix with the dimensions of $N_g \times N_g$. The entries are the joint probability density of pairs of gray levels that occur at pairs of points separated by the displacement vector d. Suppose $pd_{(I, j)}$ denotes the cardinality of the set of pairs of points that have gray level values of i and j for a movement vector $d = (d_x, d_y)$

$$p_d(i,j) = \left| \left\{ \begin{matrix} ((r,S),(r+dx,s+dy)): \\ I(r,S) = i, I(r+dx,s+dy) \end{matrix} \right\} \right| \qquad (9)$$

The offset of the parameters is shown in Fig (1).
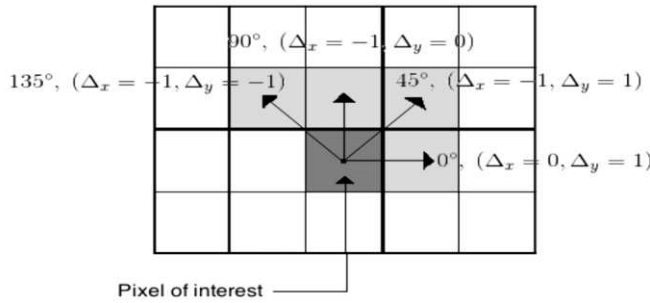
Fig 1. Illustration of the parameters Δx and Δy of the offset.

Haralick Features define the correlation in the intensity of pixels next to each other in universe [12]. Haralick suggested fourteen measures of textural features that are derivatives from the co-occurrence matrix which is a well-known statistical method for texture feature extraction. It contains data about how image intensities in pixels with a certain position in relative to each other occur together. Texture is one of the greatest important defining characteristics of an image. The grey level co-occurrence matrix is the two-dimensional matrix of joint probabilities $p(i,j)$ between pairs of pixels parted by a distance 'd' in a direction 'r'.

The second order image histogram denoted to as the Grey Level Co-occurrence Matrix (GLCM) of an image proposes a greater data about the inter-pixel relation, periodicity and spatial grey level dependencies. This matrix is a source of fourteen texture descriptors. Fig (1) illustrates the example of a GLCM along with the meaning of the offset parameters Δx and Δy[13].

These features are statistical calculations (entropy and some of variance) and (texture) to describe the image by values or specific features. Depending on the Haralick function, it extracts these features from the image.

Entropy Measures: the randomness of a gray-level distribution that is estimated to be high if the gray levels are spread randomly throughout the image [7, 14].

$$Entropy = -\sum_i^M \sum_j^N p[i,j] \log[i,j] \qquad (10)$$

Energy (Angular Second Moment): Measures the number of recurrent pairs, which is estimated to be high if the occurrence of repeated pixel sets is high.

$$Energy = \sum_i^M \sum_j^N p^2[i,j] \qquad (11)$$

Contrast: Measures the local contrast of an image. The contrast is expected to be low if the gray levels of each pixel pair are similar.

$$Contrast = \sum_i^M \sum_j^N (i-j)^2 \, p[i,j] \qquad (12)$$

Homogeneity: Measures the local homogeneity of a pixel pair, which is estimated to be great if the gray levels of each pixel pair are like.

$$Homogeneity = \sum_i^M \sum_j^N \frac{p[i,j]}{1+|i-j|} \qquad (13)$$

Sum Mean (Mean): Provides the mean of the gray levels in the image which is expected to be great if the sum of the gray levels of the image is high.

$$SumMean = \sum_i^M \sum_j^N (ip[i,j] + j[i,p]) \qquad (14)$$

Variance: Variance expresses us how spread out the distribution of gray levels that is estimated to be large if the gray levels of the image are spread out greatly.

$$Variance = \frac{1}{2}\sum_i^M \sum_j^N (i-\mu)^2 p[i,j] + (j-\mu)^2 p[i,j] \quad (15)$$

Correlation provides the relation between the two pixels in the pixel pair. They are expected to be high if the gray levels of the pixel pairs are highly correlated.

$$Correlation = \sum_i^M \sum_j^N \frac{(i-\mu)(j-\mu)p[i,j]}{\sigma^2} \qquad (16)$$

Maximum Probability (MP): Results in the pixel pair that is greatest predominant in the image that is estimated to be high if the occurrence of the maximum predominant pixel pair is high.

$$MP = M_{i,j}^{M,N} ax \, p[i,j] \qquad (17)$$

Inverse Difference Moment (IDM): expresses the smoothness of the image, like homogeneity which is estimated to be high if the gray levels of the pixel pairs are similar.

$$IDM = \sum_i^M \sum_j^N \frac{p[i,j]}{|i-j|^k} \; i \neq j \qquad (18)$$

Cluster Tendency (CT): Measures the combination of pixels that have alike gray level values [15].

$$CT = \sum_i^M \sum_j^N (i+j-2\mu)^k \, p[i,j] \qquad (19)$$

2.3 Zernike Moments Algorithm

Originally, they are polynomials Zernike polynomials known in mathematics and these polynomials are discriminated between them so that if applied to the image, they provide the program with features that aren't intertwined with each other and aren't affected by rotation and don't change the size [15].

The zernike moments is defined as the mapping of an image onto a set of complex Zernike polynomials. Therefore, Zernike polynomials are orthogonal to each other [16]. It used in visual character recognition for their properties such as heftiness to noise, expression efficiency and multi-level representation for describing the shapes of designs. Zernike moments are invariant to global rotation, translation, and scaling. It used in many uses of image analysis, reconstruction, and recognition [16].

Equation (20) is a general formula describing those moments for binary images. The variable n in this equation represents the order of a Zernike moment. When all Zernike moments are calculated up to order N, the Zernike moments are used where $n \leq N$, where N is the most order of Zernike moments.

$$z_{nm} = \frac{n+1}{\pi} \sum_x \sum_y f(x,y) R_{nm}(p) \exp(-im\theta) \qquad (20)$$

Where:

$$R_{nm}(p) = \sum_{s=0}^{n-\frac{|m|}{2}} (-1)^s \frac{(n-s)!}{s!\left(\frac{n+|m|}{2}-s\right)!\left(\frac{n-|m|}{2}-s\right)!} * p^{n-2s} \tag{21}$$

$f(x, y)$: is a digital image function.

$R_{Nm}(p)$: is the Zernike radial polynomial.

n: represents a positive integer or zero, m is non-zero integer subject to constraints is even.

p: represents the length of the vector from origin to (x, y).

θ: represents the angle between vector p and x axis in the counterclockwise direction.

As Zernike moments can be calculated to whatever order, this makes them more suitable for pattern recognition. The higher number of moments gives better performance [7, 13] the advantages of Zernike moments.

• Simple rotation invariance

• Higher accuracy for detailed shapes

• Orthogonal

• Less information redundancy

• Much better at image reconstruction (vs. normal moments).

3. Preparing Data Set and Feature Extraction

The object recognizes system contains four steps: pre-processing and segmentation image, object localization, feature extraction and classification of the image to recognize the object [17].

3.1 Processing the Image

Segmentation of Image and localization of object, it is necessary to divide the image into regions and destination objects which is contained and make use of local characteristics. Image segmentation splits a digital image into multiple segments. It is very important in many applications for any image. So far, efforts and challenges are still being made to improve segmentation techniques.

With the improvement of computer processing abilities, there are several segmentation techniques of an image: threshold, growing region, active contours, and level sets [18].

After dividing the original image into several separate regions that correspond to objects in a section, it can extract the feature vector from each region and can be measured as a representation of an object in the complete image.

3.2 Feature Extraction

Feature extraction refers to the extraction of useful information from raw data so they are suitable for the classification process. The feature extraction step is characterized by a series of input patterns.

Feature extraction methods can be divided into two directions: structural features and statistical features. The structural features deal with local data. Object change in environmental conditions is the main problem of the structural features [7, 19].

The steps for extraction features of an image are illustrated in the flowchart in Fig (2) which shows the AI algorithm for recognizing the object in the image [20].
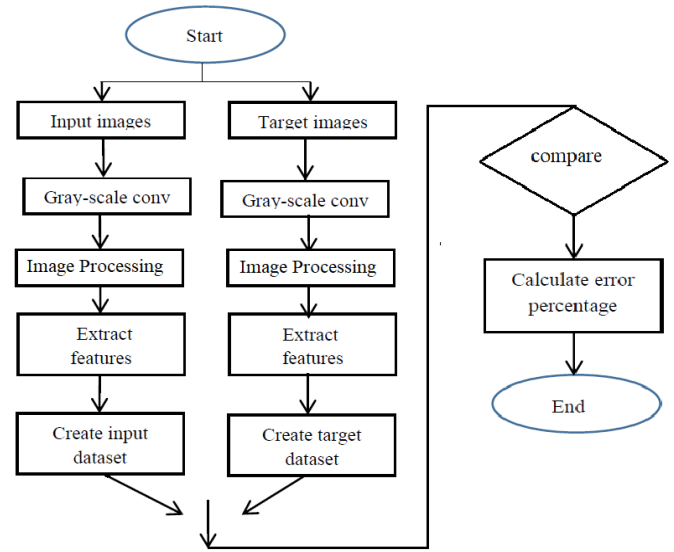

Fig 2. Structural Extract Features

3 Image Classifications and Object Recognition

Using Euclidean distance method for Zernike moments to compare between the vector moments for (features of the object) the sample of image which are trained and stored in the library with a vector of income determinations of each image sample stored in a complete public library, the following relation as shown in equation (22).

$$D_I = \sqrt{\sum_{j=1}^{m}\left(\phi_j - \phi_j^I\right)^2} \tag{22}$$

Where:

I - Represents a line guide (image guide) in the library.

D_I - The distance between the vector and the image vector moments income reference ϕ_j^I stored in line I of the library [21].

Φj - Moments image income vector.

When the distance is small and close to zero, D_I identifies the object in the image compared to the object in the photos that were previously stored in the digital library for classification and identification. Whenever the distance is of great value, the error is large and it reduces the ability to identify the object in the image.

4. Comparison Analysis and Results Discussion

Applying AI algorithms using Matlab to recognize the objects in the image and distinguish the object from the background through preparing a digital library of images have been set, where the number of images is 72 pictures. Each algorithm was studied alone on test samples for the object on which the algorithm was trained. The result for each algorithm is summarized in the following:

4.1 The Result for Zernike Moments Algorithm

Applying Zernike moments Algorithm for test samples has shown that the error rate was high and the algorithm has recognized 50% of the object as shown in Fig (3). The program also gives negative results of the objects which isn't trained. Therefore, the performance of this algorithm is weak with the process of identification. The mathematical method applied to the algorithm of Zernike moments found that the algorithm recognizes the images

that are trained with a match rate (80-93) % of the samples which were trained in the program as shown in Fig (4).

The Result of Haralick Features Algorithm

The algorithm is training process for images in library and a matrix of statistical information is created for each image by using Matlab program. The algorithm is applied for the test images that isn't trained by the algorithm. The program gave positive results and the object is recognized that shows in Fig (5). Whereas, the error rate was small, (1-2.5) %, and this result is an acceptable. The results illustrate that this algorithm is excellent for work and can be used in distinguishing and recognizing objects with high accuracy.

4.3 The Result of Hue Moment algorithm

This algorithm gives excellent results and is identical to a large amount with the samples of objects which were trained when applied to the test samples. Images are the same of the object that algorithm has not trained for the test sample. Note that in Fig (6) the error rate doesn't exceed 0.1%, thus there is a high accuracy of recognition. It is very high accuracy compared with the previous algorithms and is identical to the error values of the samples on which the algorithm is trained Applying the Hue moment algorithm to the test images, the results are small error rate Notice that this algorithm is contributed in increasing the accuracy of the recognition objects significantly.

When is applying the algorithm for the images of other objects which that isn't trained on, the error rate is zero and the algorithm didn't recognize for the object within the programmatic conditions and it gave positive results as shown in Fig (7). The results illustrate that this algorithm is excellent for work and can be used in distinguishing and recognizing of objects with high accuracy.
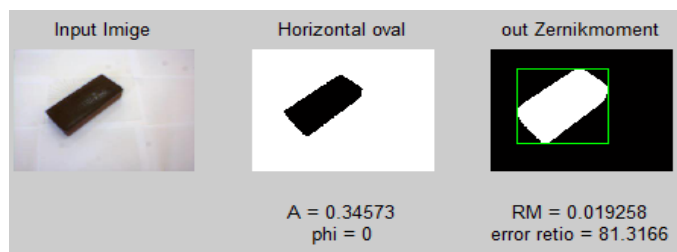

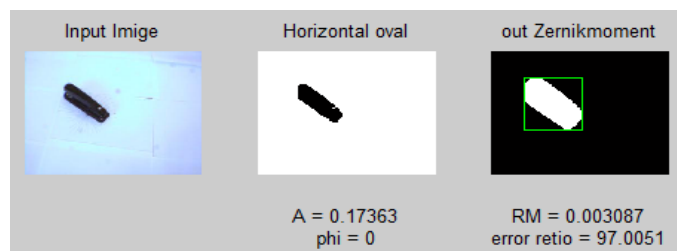Fig.3. Error Rate in Zernike Moments Algorithm

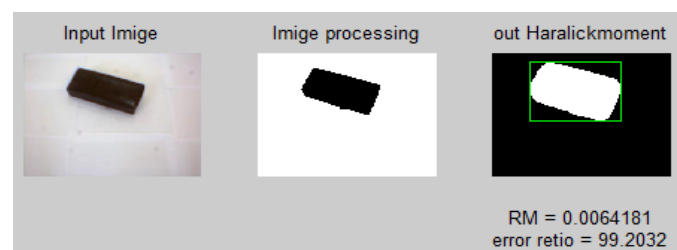**Fig 4.** The Error Rate in Zernike moments for the Image Test

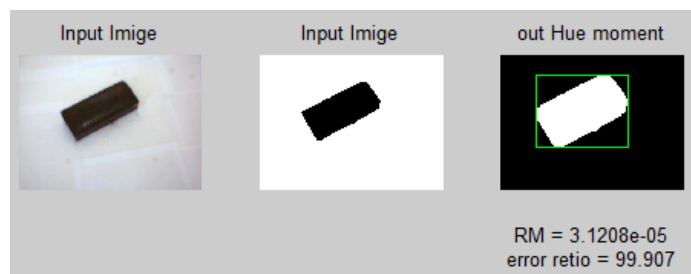**Fig 5.** Recognizing Process by Haralick features.

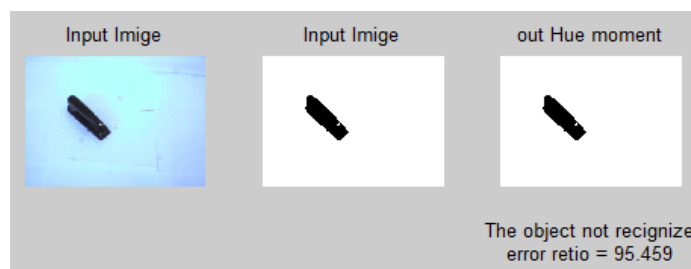**Fig 6.** Recognizing process by the Hue moments algorithm

Fig 7. Recognition process for the test samples in Hue moments

5. Conclusion and Recommendation

AI algorithms create a digital library containing samples and information for training to work and distinguish the objects. The digital images were processed using digital image processing technology that includes noise removal, filtering, image thresholding and increasing the sharpness of the object details until it becomes ready to be recognized by the proposed algorithm. The algorithm of Zernike moment performance of was weak and the error rate is large. It wasn't effective. After training the Haralick Features algorithm on the proposed images, the efficiency of algorithm was high and gave an excellent result with few errors in the process of the recognition. Hue moment algorithm has given the highest efficiency in recognizing objects. When applied on test samples the error rate doesn't exceed 0.1% and almost equal to zero, but when applied to images that haven't been trained, the error rate was raised to 5% which is also acceptable.

The previous discussed algorithms can be developed and integrated to work with artificial neural networks.

REFERENCES

[1] Radu Tudor Ionescu, et al. Patch Autocorrelation Features: A translation and rotation invariant approach for image classification, Artificial Intelligence Review, ISSN 0269-2821.2016.

[2] Alba.L. Rossell. and Josep.R.Casas. 3D reconstruction and recognition of the object using a Kinect camera, Polytechnic University of Catalonia, 2016.

[3] Mohd Hana. And Ahmad Hijazi. Image Classification: A Study in Age Related Macular Degeneration Screening, University of Liverpool, 2012.

[4] Bhaskara Rao, Vara Prasad, and Ch.p. Kumar. Feature Extraction Using Zernike Moments, IJLTET, Vol. 2 Issue, 2013.

[5] Thawar Arif, et al. object classification via geometrical Zernike and Legendre, JATIT, Vol 7, No 1, 2014

[6] Mohd.W. et al. Image Extraction using Geometric and Zernike Moment Invariants. University Malaysia Perlis, Kangar, Perlis, 2016.

[7] Habib Mahi, Hadria Isabaten. and Chahira Serief. Zernike Moments and SVM for Shape Classification in Very High Resolution Satellite Images, The International Arab Journal of Information Technology, Vol 11, No1, 2014.

[8] Simon Xinmeng Liao. Image Analysis by Moments, University of Manitoba, Canada. 1993.

[9] Hsin-yi Tsai. et al. GPU-Accelerated Features Extraction from Magnetic Resonance Images, IEEE, ACCESS, 2017.

[10] Maofu Liu, Yanxiang He & Bin Ye. Image zernike moments shape feature evaluation based on image reconstruction, Geo-spatial Information Science, Vol 10, Issue 3, 2007.

[11] Marouf. Hoda. and Karim Faez. Zerinke moment -based feature extraction for facial recognition of identical twins, IJCSEIT, Vol3, 2013.

[12] P. Mohanaiah, P. Sathyanarayanan, L. and GuruKumar. Image texture feature extraction using GLCM approach, International Journal of Scientific and Research Publications, V 3, Issue 5,2013.

[13] Johan Wiebe. Gender identification from video sequences. Chalmers University of Technology Gothenburg, Sweden, 2014.

[14] Dahiya Susheela, P. K. Garg. and Mahesh. JatBuilding Extraction from High Resolution Satellite Images, Malaviya National Institute of Technology Jaipur, 2014.

[15] Tanzia Ahmed, Bir.B. Roy, and Tanvir Rahman. Drone Detection by Neural Network using SURF Feature and GLCM Method. BRAC University, 2017.

[16] Karbhari.V. Kale, Prapti.D. and Deshmukh. Zernike Moment Feature Extraction for Handwritten Devanagari (Marathi) Compound Character Recognition, IJARAI, Vol 3, No, 2014.

[17] Oujaoura Mustapha, Brahim Minaoui. and Mohammed Fakir. Image Annotation by Moments, GCSR Vol. 1, @ Science Gate Publishing, 2014.

[18] Gottipati. S. Babu. Moving object detection using Matlab, IJERT, Issue 6, 2012

[19] Faleh H. Mahmood, and W.A. Abbas. Texture Features Analysis using Gray Level Co-occurrence Matrix for Abnormality Detection in Chest CT Images, Iraqi Journal of Scienc, 2013.

[20] Arati S. Kurani, et al. Co-occurrence matrices for volumetric data. Intelligent Multimedia Processing Laboratory, School of Computer Science, DePaul University, Chicago, Illinois, 2012

[21] Cecilia.D. Ruberto, Lorenzo Putzu, and Giuseppe Rodriguez. Fast and accurate computation of orthogonal moments for texture analysis. University of Cagliari, piazza Italy, 2018.