Identification of Similar Looking Bulk Split Grams using GLCM and CGLCM Texture Features

Pushpalatha K. R.
Assistant Professor
Dept. of MCA
Sri Siddhartha Institute of Technology,Tumkur
Karnataka, India.

Asha Gowda Karegowda
Associate Professor
Dept. of MCA
Siddaganga Institute of Technology, Tumkur
Karnataka, India

D. Ramesh
Professor
Dept. of MCA
Sri Siddhartha Institute of Technology, Tumkur
Karnataka, India

ABSTRACT
Content based image retrieval (CBIR) is an automated way to retrieve images based on the visual content or image features itself. Visual inspection of food type is tiresome and time consuming task. This paper presents the retrieval of similar looking bulk split gram images using Grey Level Co-occurrence Matrix (GLCM) and Color Grey Level Co-occurrence Matrix (CGLCM) texture features. Texture feature matching procedure is based on three distance measures namely, Euclidean distance, Canberra distance and City block distance. The performance of a retrieved image is measured in terms of Precision. Experimental results show that the CGLCM provides better retrieving result than GLCM.

Keywords
CBIR, GLCM, CGLCM, Euclidean Distance, Canberra Distance, City Block Distance, Precision.

1. INTRODUCTION
CBIR is effective image retrieval method widely used both in academia and industries. In agriculture industry the food grain type and the quality are identified by human inspection from long back. At present the food grain type and quality are rapidly assessed through visual inspection by human inspectors. This process is tiresome and takes a lot of time to do. Human inspection decision making affect by external influence such as tiredness, vengeance, prejudice etc. These tasks require automatic recognition system to identify the food grain images. Hence Machine Vision System is an alternative to this manual practices [1], [2].

Anami et.al and visen et al. have used an Artificial Neural Network (ANN) approach to identify and classify the bulk food grain samples [3], [4], [5]. Anami et al. have tried to use color feature extraction for boiled food grains [6]. Neelamma K Patil et al. have used color feature extraction for classification of different food grains [7], [8], [9]. Anami et al. used color and texture feature extraction for identification and classification of normal and affected fruits, cereals and grains[10]. Dayanand Savakar worked on an identification and classification of bulk grains and fruits using ANN [11], [12]. P.K. Mallick et.al and Ksh.Rober Singh et.al carried out the comparative study on ANN, Back Propagation Neural Network (BPNN) classifier with Support Vector Machine (SVM), K-Nearest Neighbor(KNN) and Naïve Bayes Classifier[13],[14]. Jagadeesh .D.Pujari and S.J.Mosavi Rad et.al. proposed reduced future set selection based method for normal and affected food grains [15],[16],[17]. Iman Golpour et.al and S.Majumdar et.al. suggested STEPDISC analysis method for feature selection [18],[19]. Anami et.al projected feature extraction methods for identification and classification of food grains, fruits and flowers [20]. Anami etl.al. recommended edge texture features based methodology for recognition of variety of bulk paddies [21].

Authors classified food grains into split grams and whole food grains. In this paper authors worked on split gram images. Figure.1 shows the block diagram of CBIR process. CBIR mainly consists of two steps: feature extraction and feature matching. In feature extraction the features are extracted from split gram image. The extraction process is done on both query split gram images and database split gram images using GLCM and CGLCM. Feature matching involves similarity between features of images compared and computed using Euclidean, Canberra and City block distance. The present work involves retrieval of 8 different classes of similar looking bulk split grams. GLCM and CGLCM are used to extract the texture features. Precision is used to measure the performance of CBIR for different split grams. Feature matching procedure is based on three distance measures namely, Euclidean distance, Canberra distance and City block distance.

Rest of this paper is as follows, Section two describes the proposed work and methodology. Section three and four briefs feature extraction methods and feature matching respectively. Section five describes performance criteria. Results and discussion in section six followed by conclusion in section seven.

2. PROPOSED WORK AND METHODOLOGY
This paper presents the study on identification of spilt grams using GLCM and CGLCM texture features. In this paper, the experiment was carried on 8 different types of split grams. The split grams considered in this work includes, Split pigeon peas of type1, Split red lentil , Split pigeon peas of type 2, Split bengal gram , Split black gram ,Roasted split bengal gram, Split fava gram, Split green gram are shown in the Figure 2. The original images used for work was captured under natural light and resized to 250X250. Features are extracted by GLCM and CGLCM texture feature. Euclidean distance, Canberra and City block distance has been used for feature matching. Precision used as performance evaluation method.

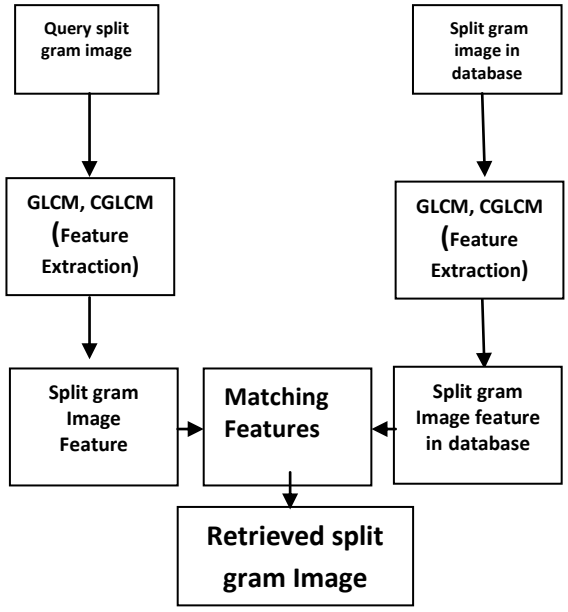

Fig. 1: Block Diagram of proposed work

Fig. 2: Images of Similar looking bulk split gram Samples (a)Split pigeon peas type1 (b) Split red lentil type1 (c) Split pigeon peas type2 (d) Split bengal gram (e) Split blackgram (f) Roasted split bengal gram (g) Split fava gram (h) Split green gram

3. TEXTURE FEATURE EXTRACTION

Texture plays an important role in many machine vision tasks such as surface inspection, scene classification and surface orientation and shape determination. In order to capture the spatial dependence of gray-level values which contribute to the perception of texture, a two-dimensional dependence matrix known as gray-level co-occurrence matrix is extensively used in texture feature extraction.

3.1 Grey level co-occurrence matrix (GLCM)

Some of the split grams are similar in color but exhibit different texture features. We adopted co-occurrence matrix to obtain texture features. The co-occurrence matrix method of texture description is based on the repeated occurrence of some gray-level configuration in the texture. The part of textured image to be analyzed is an M*N rectangular window. Some gray-level configuration may be described by a matrix of relative frequencies $P_{\alpha,d}(i,j)$,describing how frequently two pixels with gray-levels i,j appear in the window separated by a distance d in the direction α. The co-occurance matrix procedure is given in Algorithm 1.

Algorithm 1: Calculation of co-occurrence matrix from the image f (i,j)

Input: gray level image of size (M*N)

Output: Co-occurrence matrix $P_{\alpha,d}(i,j)$ for d=1 in the direction α.

Start

Step 1: Assign Pα, d (i, j) =0 for all x, y Ɛ [0, L] where L is the maximum gray level.

Step 2: For all pixels (i_1, j1) in the image determine (i_2, j2) and perform

$$P_{\alpha,d}[f(i_1,j_1),f(i_2,j_2)]=P_{\alpha,d}[f(i_1,j_1),f(i_2,j_2)]+1$$

Stop.

The features extracted from GLCM method are as follows: Contrast, Correlation, Energy, Entropy, Homogeneity, Dissimilarity, Inverse difference, Cluster Shade, Cluster Prominence, Autocorrelation, Maximum probability, Sum of Squares, Sum Average, Sum Variance, Sum Entropy, Difference variance, Difference entropy, Information measures of correlation (1), Information measures of correlation (2), Maximal correlation coefficient, Inverse difference normalized (INN), Inverse difference moment normalized (IDN) .

Haralick extracted 22 GLCM texture features are briefed as follows [22] :**Contrast :**

$$f_2=\sum_{n=0}^{\theta}n^2\left[\sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1}p_{d,\theta}(i,j)\right],n=|i-j|$$

When i and j are equal, the cell is on the diagonal and i - j = 0. These values represent pixels entirely similar to their neighbor, so they are given a weight of 0. If I and j differ by 1, there is a small contrast, and the weight is 1. If i and j differ by 2, the contrast is increasing and the weight is 4. The weights continue to increase exponentially as (i - j) increases.

Entropy: $$f_3=\sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1}P_{d,\theta}(i,j)\log(P_{d,\theta}(i,j))$$

Variance: $$f_4=\sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1}P_{d,\theta}(i,\mu)^2P_{d,\theta}(i,j))$$

Correlation:

$$f_5=\sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1}P_{d,\theta}(i,j)\frac{(i-\mu_x)(j-\mu_y)}{\sigma_x\sigma_y}$$

Energy: $$f_6=\sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1}P_{d,\theta}(i,j)^2$$

Homogeneity or Inverse difference moment:

$$f_7=\sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1}\frac{1}{1+(i-j)^2}P_{d,\theta}(i,j)$$

.**Sum Average:** $f_8 = \sum_{i=0}^{2(N_g-1)} i.p_{x+y}(i)$

Where $p_{x+y}(k) = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} P_{d,\theta}(i, j)$

$, k = i + j = \{0,1,2,...2(N_g - 10)\}$

Sum variance: $f_9 = \sum_{i=0}^{2(N_g-1)} (i - f_8)^2 p_{x+y}(i)$

Sum entropy: $f_{10} = \sum_{i=0}^{2(N_g-1)} p_{x+y}(i) \log p_{x+y}(i)$

Difference variance: $f_{11} = \sum_{i=0}^{2(N_g-1)} (i - f_{11}^{'})^2 p_{x-y}(i)$

Where, $p_{x+y}(k) = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} P_{d,\theta}(i, j)$

$, k = i + j = \{0,1,2,...2(N_g - 10)\}$

$f_{11}^{'} = \sum_{i=0}^{(N_g-1)} i.p_{x-y}(i)$

Difference entropy: $f_{12} = \sum_{i=0}^{(N_g-1)} p_{(x-y)}(i) \log p_{x-y}(i)$

Information measure of correlation 1:

$$f_{13} = \frac{HXY - HXY_1}{\max(HX, HY)}$$

Information measure of correlation 2:

$$f_{14} = (1 - \exp[-2(HXY_2 - HXY])^{1/2}$$

Where $HX = -\sum_{i=0}^{N_g-1} p_x(i) \log(p_x(i))$

$$HY = -\sum_{i=0}^{N_g-1} p_y(i) \log(p_y(i))$$

$$HXY_1 = -\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} p_{d,\theta}(i, j) \log(p_x(i) p_y(j))$$

$$HXY_2 = -\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} p_x(i) p_y(j) \log(p_x(i) p_y(j))$$

Cluster shade:

$$SHADE = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \{i + j - \mu_x - \mu_y\}^3 XP(i, j)$$

Cluster prominence:

$$PROM = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \{i + j - \mu_x - \mu_y\}^4 XP(i, j)$$

3.2 Color grey level co-occurrence matrix (CGLCM)

CGLCM is the special type of texture feature extraction. A simple extension of GLCM technique is Color GLCM for color images. In color space the texture analysis techniques and their statistical features are computed for each band like red, green, blue. Comparisons can then be done between similar bands from two different images for classification. In this case, feature vector is computed directly from 3D RGB color space, where for distance d=1 the cube of size 3x3x3 is created. By moving this cube in an image, three GLCM matrices are computed for every channel. 22 statistical features are extracted from GLCM. The color information is captured by applying GLCM to each of the color channels ***r,g,b*** individually. For each color channel 22 features are extracted resulting in 66 features. The example of computing CGLCM for G (Green channel is described by equations [23], [24].

$$GLCM_{i,j}(G) = \sum_{m=-1}^{1} \sum_{n=-1}^{1} relation(img(i, j, 2) | img(i + m, j + n, 2))$$

$$GLCM_{i,j}(GR) = \sum_{m=-1}^{1} \sum_{n=-1}^{1} relation(img(i, j, 2) | img(i + m, j + n, 1))$$

$$GLCM_{i,j}(GB) = \sum_{m=-1}^{1} \sum_{n=-1}^{1} relation(img(i, j, 2) | img(i + m, j + n, 3))$$

Where *img* is an split gram image represented by RGB color space, *G* is representing the *G* channel of RGB, and GR and *GB* are relations of *G* channel to *R* and *G* channel.

4. FEATURE MATCHING

A variety of metrics have been proposed to calculate the distance between the query split gram image and split gram image in database. Distance between two split gram images has to be calculated to find if there is any match or not. Distance will help us in finding degree of matching for the entire split gram data base. There are different distances available which are used and compared the performance of all these distances. In this proposed work three methods are used namely Euclidian Canberra and City block Distance Measurements [2].

Euclidean Distance

$$d_{ED}(i, j) = \sqrt{\sum_{i=1}^{m} (x_{ij} - x_{jk})^2}$$

Where x_{ij} and x_{jk} are the feature vectors of split gram image and query split gram image respectively.

City Block Distance

This is always greater than or equal to zero. The measurement would be zero for identical points and high for points that show little similarity.

$$d_{CB}(i,j) = \sum_{k=1}^{n} \left| x_{ij} - x_{jk} \right|$$

Canberra Distance
The sum of series of fraction differences between coordinates of a pair of objects.

$$d_{CD}(i,j) = \sum_{k=1}^{n} \frac{\left| x_{ik} - x_{jk} \right|}{\left| x_{ik} \right| + \left| x_{jk} \right|}$$

5. PERFORMANCE EVALUATION

The performance of a CBIR for split grams is measured in terms of its precision. Precision measures the ability of the system to retrieve only models that are relevant and is computed as follows:

$$\text{Precision (p)} = \frac{c_p}{p}$$

Where, p the number of retrievals, c_p is the number of relevant matches among all the p retrievals,

6. RESULTS AND DISCUSSION

Texture Features are extracted using GLCM and CGLCM with Euclidean, Canberra and City block distance. The performance of a retrieval system is measured in terms of precision.

Performance of split gram images measured in terms of precision measures using features extracted by GLCM, CGLCM for different k values for Euclidean distance, Canberra and City block is shown in Table 1. Precision for the 8 types of split grams was 100% for all the three distance measures with K=5, 10, 12. Table 2 presents the average precision for K=5,10 and 12 for all the 8 split gram types using three distance measures using Euclidean, Canberra and City block distance. Figure.3 depicts the comparative results of performance of proposed CBIR system using Euclidean distance for 8 different types of split grams. Figure.4 depicts the comparative results of performance of proposed CBIR system using Canberra distance for 8 different types of split grams. Figure.5 depicts the comparative results of performance of proposed CBIR system using City block distance for 8 different types of split grams. Figure.6 depicts the comparative average results of proposed CBIR for Euclidean, Canberra and City block distance. The average precision (for K=5,10,12) of proposed CBIR for split grams using GLCM with Euclidean distance was minimum with 0.4 for split red lentil, split bengal gram and split black gram. Precision was better for split pigeon, peas type 2 and split green gram for Euclidean, Canberra and City distance measure.

7. CONCLUSION

CBIR for split grams using texture features can overcome the time consuming and tiresome task of visualization of food types by food inspectors. It was observed that the performance of CBIR for split gram images using CGLCM features was almost consistent for different value of k and three distance measures namely: Euclidean, Canberra and City block distance. Experimental results show that the CGLCM provides better retrieving result compared to GLCM for all the 8 types of split grams. Authors would like to carry out similar kind of work on bulk similar looking whole food grains.

8. REFERENCES

[1] Asha Gowda Karegowda, Bharathi, "Enhancing CBIR Performance using Evolutionary Algorithms Assisted Significant Feature Selection", A Filter Approach International Journal of Applied Research on Information Technology and Computing,Vol. 7,No.1,Page:53-59, April 2016.

[2] Asha Gowda Karegowda, Pushpalatha.K.R, Chaitra.M," Color Histogram Based Image Retrieval –A survey", International Journal of Advanced Research in computer science, Vol 4, Page: 119-126, Nov-Dec 2013.

[3] N.S. Visen, J. Paliwal, D.S. Jayas,N.D.G. White, "Image analysis of bulk grain samples using neural networks", Journal of Canadian Biosystems Engineering, Vol. 46, Page:7.11-7.15, 2004.

[4] Basavaraj .S. Anami, Vishwanath.C.Burkpalli, "Texture based Identification and Classification of Bulk Sugary Food Objects", International Conference on Computer Science and Engineering (ICGST), Vol.9, Issue 4, Page: 9-14, 2009.

[5] Basavaraj S. Anami Naveen N. M.2 ,N. G. Hanamaratti ,"Behavior of HSI ColorCo-Occurrence Features in Variety Recognition from Bulk Paddy Grain Image Samples", International Journal of Signal Processing and Pattern Recognition,Vol. 8, No. 4, Page:19-30, 2015.

[6] Basavaraj S. Anami, Vishwanath C burkpalli, "Color Based Recognition and Estimation of Temperature Levels of Images of Boiled Food Grains", International Journal of Computer Applications, Vol.1 No.14,Page:98-103, 2010.

[7] Neelamma K. Patil Ravi M. Yadahalli Jagadeesh Pujari, "Comparison between HSV and YCbCr Color Model Color-Texture based Classification of the Food Grains", International Journal of Computer Applications, Vol.34, No.4, Page: 51-57, 2011.

[8] Neelamma K. Patil, Virendra S. Malemath, Ravi M. Yadahalli, "Color and Texture Based Identification and Classification of food Grains using different Color Models and Haralick features", International Journal on Computer Science and Engieering (IJCSE),Vol.3, No.12,Page:3669-3680,Dec 2011.

[9] Neelamma K. Patil, Ravi M. Yadahalli, Jagadeesh Pujari, "The Effect of Block Size, Training set and K-value in the Classification of Food Grains using L*a*b color model by combining Color and Texture Information", Global Journal of Computer Application and Technology, Vol.2, No.1, Page: 884-889, 2012.

[10] Basvaraj .S. Anami, Jagadeesh D.Pujari,Rajesh Yakkundimath,"Identification and Classification of Normal and Affected Agriculture/horticulture Produce Based on Combined Color and Texture Feature Extraction", International Journal of Computer Applications in Engineering Sciences, Vol.1, Issue 3, Page:356-360, 2011.

[11] Dayanand Savakar,"Recognition and Classification of Similar Looking Food Grain Images using Artificial Neural Networks", Journal of Applied Computer Science and Mathematics, Vol.6, No. 13, Page: 61-65, 2012.

[12] Dayanand Savakar, "Identification and Classification of Bulk FruitsImages using Artificial Neural Networks",

International Journal of Engineering and Innovative Technology (IJEIT), Vol. 1, Issue 3, Page: 36-40, 2012.

[13] P.K. Mallick, S.Rout, "Identification and Classification of Similar Looking Food Grains", International Journal on Advanced Computer Theory and Engineering (IJACTE), Vol. 1, Issue 2, Page: 139-144, 2013.

[14] K s h .Robert Singh, Saurabh Chaudhury, "An Efficient Technique for Rice Grain Classification using Back Propagation Neural Network and Wavelet Decomposition", IET Computer Vision 2016.

[15] S. J. Mousavi Rad, F. Akhlaghian Tab, K. Mollazade, "Application of Imperialist Competitive Algorithm for Feature Selection: A Case Study on Bulk Rice Classification", International Journal of Computer Applications, Vol.40, No.16, Page: 41-48, 2012.

[16] Jagadeesh. D. Pujari, Rajesh. Yakkundimath, A. S. Byadgi, "Reduced Color and Texture features based Identification and Classification of Affected and Normal fruits' images", International Journal of Agricultural and Food Science,Vol.3, No.3, Page:119-127, 2013.

[17] Jagadeesh D.Pujari,Rajesh Yakkundimath, Abdulmunaf S. Byadgi,"Recognition and Classification of Normal and Affected Agriculture Produce using Reduced Color and Texture Features", International Journal of Computer Applications, Vol.93, No.11, Page:17-24, 2014.

[18] Majumdar S., Jayas D.S "Classification of bulk samples of cereal grains using machine vision", Journal of Agricultural Engineering Research, Vol.73, Page:35-47, 1999.

[19] Iman Golpour, Jafar Amiri Parian ,Reza Amiri Chayjan, "Identification and Classification of Bulk Paddy, Brown,and White Rice Cultivars with Colour Features Extraction using Image Analysis and Neural Network", Czech Journal of Food Science,Vol.32, No.3, Page:280-287, 2014.

[20] Basavaraj S. Anami, Dayanand G. Savakar, "Suitability of Feature Extraction Methods in Recognition and Classification of Grains,Fruits and Flowers", International Journal of Food Engineering, Vol. 7 , Issue 1, Art.9, 2011.

[21] Basavaraj Anami, Dayanand G. Savakar, "Effect of Foreign Bodies on Recognition and Classification of Bulk Food GrainsImage Samples", Journal of Applied Computer Science, Vol.3, No.6, Page:77-83, 2009.

[22] R.M. Haralick, K. Shanmugam, and I.H. Dinstein. "Textural features for image classification" IEEE Transactions on Systems, Man, and Cybernetics, Page: 237-247, 1973.

[23] Moulay A. Akhloufi, Xavier Maldague, Wael Ben Larbi,"A New Color-Texture Approach for Industrial Products Inspection", Journal Of Multimedia, Vol. 3, No. 3, Page: 44-50, July 2008.

[24] Miroslav Benco, Robert HUDEC,"Novel Method for Color Textures Features Extraction Based on GLCM", Radio Engineering, Vol. 16, No. 4, Page: 64-67, Dec 2007.

9. APPENDIX

Table 1 : Precision for identification of different split grams using different distance measures for K=5,10,12

Split Gram Image	K value	Euclidean Distance Measures				Canberra Distance Measures				City Block Distance Measures			
		GLCM		CGLCM		GLCM		CGLCM		GLCM		CGLCM	
		Match	Precision	Match	Precision	Match	Precision	Match	Precision	Match	Precision	Match	Precision
Split pigeon Peas Type1	5	2	0.40	5	1	2	0.40	3	0.6	2	0.40	5	1
	10	6	0.60	10	1	6	0.60	7	0.7	6	0.60	10	1
	12	6	0.50	12	1	6	0.50	9	0.75	6	0.50	12	1
Split red Lentil	5	2	0.40	5	1	2	0.40	5	1	2	0.40	5	1
	10	4	0.40	10	1	4	0.40	10	1	4	0.40	10	1
	12	4	0.33	12	1	4	0.33	12	1	4	0.33	12	1
Split pigeon Peas Type2	5	4	0.80	5	1	4	0.80	5	1	4	0.80	5	1
	10	8	0.80	10	1	8	0.80	10	1	8	0.80	10	1
	12	8	0.67	12	1	8	0.67	12	1	8	0.67	12	1
Split Bengal gram	5	2	0.40	5	1	2	0.40	5	1	2	0.40	5	1
	10	5	0.50	10	1	5	0.50	10	1	4	0.40	10	1
	12	5	0.42	12	1	5	0.42	12	1	4	0.33	12	1
Split Black gram	5	2	0.40	5	1	2	0.40	5	1	4	0.80	5	1
	10	4	0.40	10	1	3	0.30	10	1	9	0.90	10	1
	12	4	0.33	12	1	4	0.33	12	1	9	0.75	12	1
Roasted	5	3	0.60	5	1	3	0.60	5	1	3	0.60	5	1

Split Bengal Gram	10	5	0.50	10	1	5	0.50	10	1	5	0.50	10	1
	12	7	0.58	12	1	7	0.58	12	1	7	0.58	12	1
Split Fava gram	5	4	0.80	5	1	4	0.80	5	1	4	0.80	5	1
	10	5	0.50	10	1	7	0.70	10	1	7	0.70	10	1
	12	7	0.58	12	1	7	0.58	12	1	7	0.58	12	1
Split Green Gram	5	5	1.00	5	1	5	1.00	5	1	5	1.00	5	1
	10	8	0.80	10	1	9	0.90	10	1	8	0.80	10	1
	12	9	0.75	12	1	10	0.83	12	1	9	0.75	12	1

Table 2: Average precision of GLCM and CGLCM using different Distance Measure for K=5,10,12

Split Gram Image	Euclidean Distance Measure		Canberra Distance Measure		City block Distance Measure	
	GLCM AVG precision	CGLCM AVG precision	GLCM AVG precision	CGLCM AVG precision	GLCM AVG Precision	CGLCM AVG precision
Split pigeon Peas Type1	0.5	1	0.5	0.6	0.5	1
Split red Lentil	0.4	1	0.4	1	0.4	1
Split pigeon Peas Type2	0.8	1	0.8	1	0.7	1
Split Bengal gram	0.4	1	0.4	1	0.4	1
Split Blackgram	0.4	1	0.3	1	0.8	1
Roasted Split Bengal Gram	0.6	1	0.6	1	0.6	1
Split Fava gram	0.6	1	0.7	1	0.7	1
Split green gram	0.9	1	1	1	0.9	1

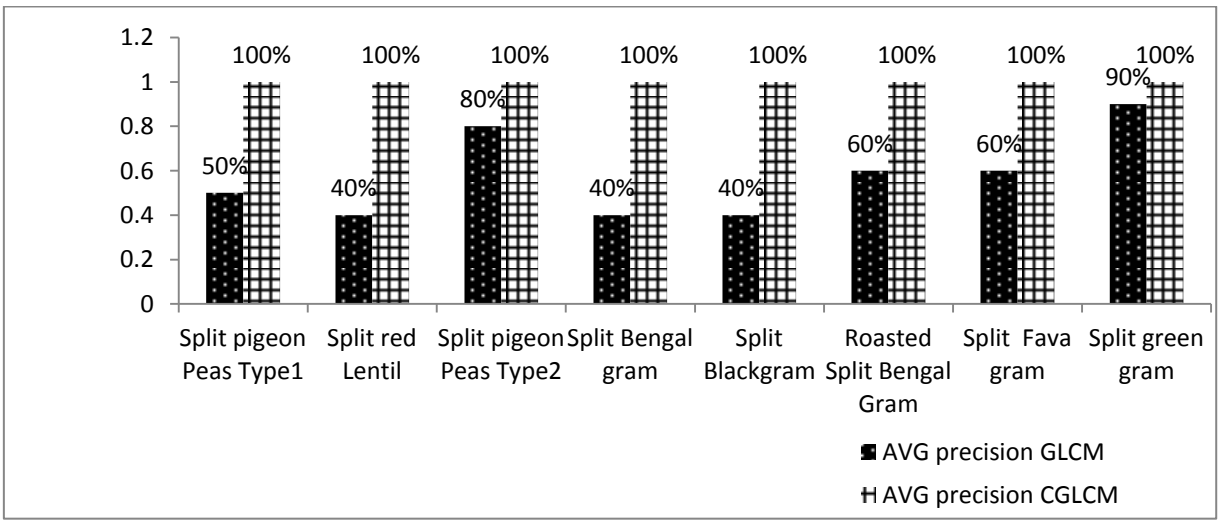

Fig. 3: Average precision for different split grams using Euclidean Distance Measure for K=5, 10, 12

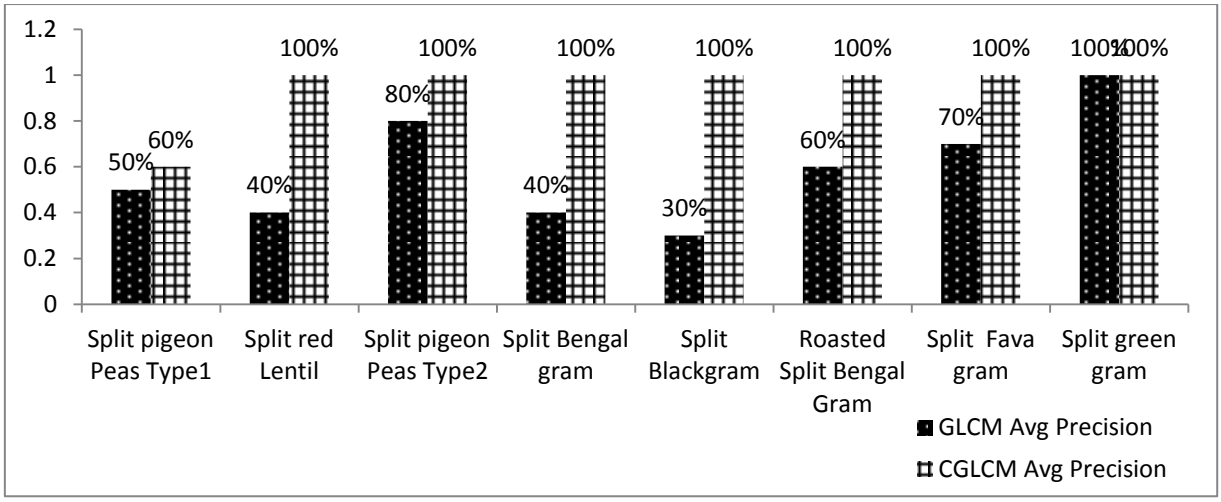

Fig. 4: Average precision for different split grams using Canberra Distance Measure for K=5, 10, 12

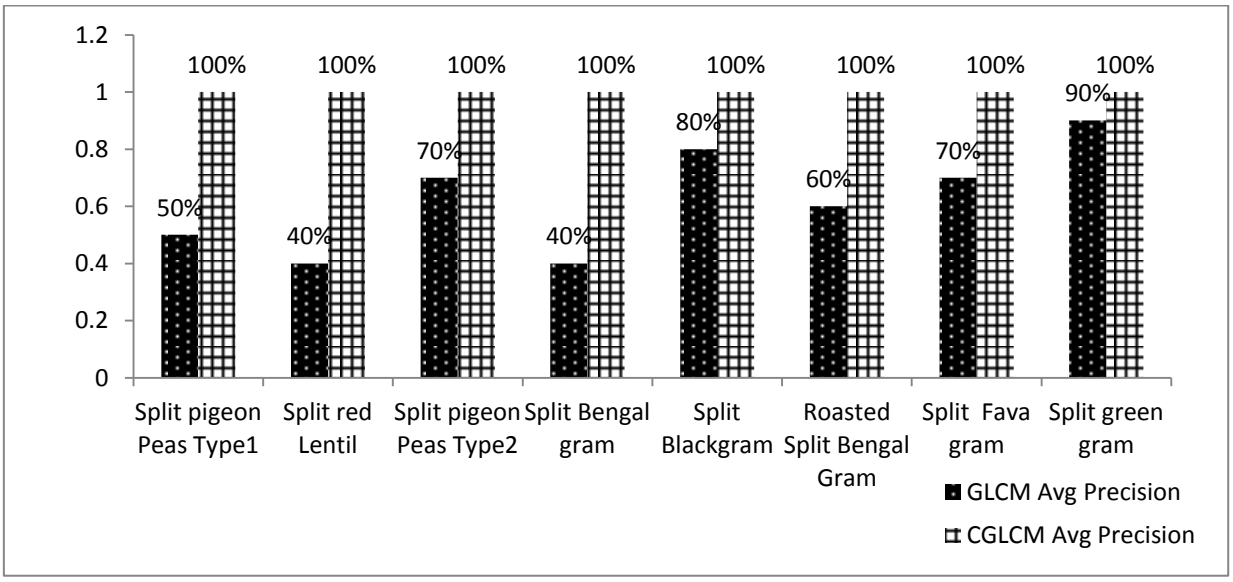

Fig. 5: Average precision in % for different split grams using City Block Distance Measure for K=5, 10, 12

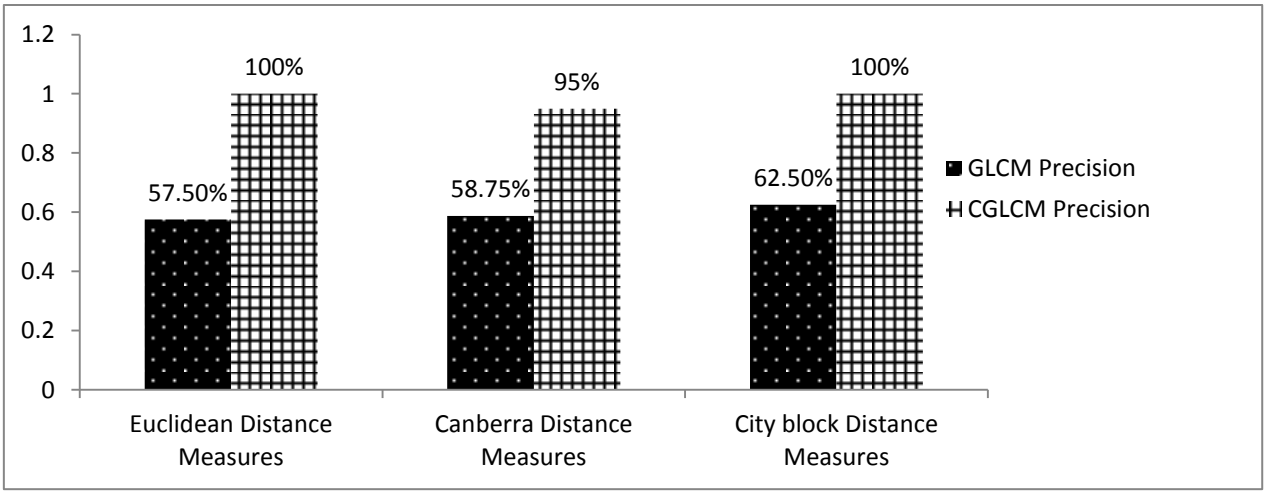

Fig. 6: Average precision for all 8 types of split gram using Euclidean, Canberra and City block Distance Measures.