Performance evaluation of CBIR techniques

Amandeep Verma, Tania Gupta

vaman71@gmail.com, tania.1830@gmail.com

Abstract— In today's epoch of Digital technology, the need of content based image retrieval is increasing day by day. Image retrieval system based on image's visual content becomes more desirable. Hence, content based image retrieval has emerged as one of the most active research areas in the past few years. There are a lot of approaches have been proposed for content based image retrieval. The literature addressed the issues involved in content based image retrieval and presents solution to them. To this end, an interface is designed for content based image retrieval and in order to key out best approach, a comparative study has been performed. Five classes namely tribal, beach, bus, building and dinosaur have been chosen from standard dataset. The CBIR techniques were compared on the basis of sensitivity, specificity, accuracy, overall accuracy etc. The result fortified that the HOG and LBP-HOG feature extraction algorithm is more efficient than Relevance Feedback and FSRM algorithms.

Keywords: CBIR, performance evaluation, relevance feedback, FSRM, HOG, LBP-HOG etc.

I. INTRODUCTION

The search for similar images in large-scale image databases has been an active research area in the last couple of years. A very promising approach is content based image retrieval. Content-based image retrieval systems were introduced to address the problems associated with text-based image retrieval [1]. Content based image retrieval is a set of techniques for retrieving semantically-relevant images from an image database based on automatically-derived image features. CBIR system basically consists of two parts i.e. feature extraction and similarity matching. During feature extraction, important features are extracted from the image and stored in a feature vector, after feature extraction distance between two feature vectors is calculated by using similarity measures.

In CBIR many techniques [2] [3] [4] [5] [6] are used to evaluate different impact on the results of retrieving images. Already a lot of work has been done since many years on CBIR techniques. To this end, the four approaches for content based image retrieval is selected for implementation. These approaches are Relevance feedback [7] and Fuzzy semantic Relevance Matrix [8]; HOG feature extraction [9] and LBP-HOG feature extraction [10] approaches. The relevance feedback technique in

CBIR is based on user interaction with the retrieval system. This technique is designed to bridge the semantic gap between low level and high level features. Histogram of Oriented Gradients (HOG) is a powerful feature descriptor used in computer vision and image processing for the purpose of object detection. The basic idea behind the HOG descriptor is that local object appearance and shape within an image can be described by distribution of intensity gradients or edge detections Local Binary Pattern is a feature descriptor like HOG. It has been used for texture classification [11]. The LBP has been widely used in different applications such as human detection, face recognition.

The paper has been organized as follows: section II discusses the related work in CBIR approaches, section III gives details implementation of CBIR techniques, section IV explains the algorithms that have been evaluated and section V discusses the results and last section VI presents the conclusion and future work.

II. RELATED WORK

There is a large array of literature associated with CBIR approaches which are presented as follows:

Zhang *et al.*[8] worked upon the semantic memory based image retrieval. The semantic memory is fulfilled using the user relevance feedback. The authors have proposed the use of semantic subspace learning (SSL) method for similarity matching between feature vectors of query image and dataset images. The proposed method compared with other distance metrics i.e. Euclidean distance and Mahalanobis distance metric. The average precision of semantic subspace learning method (SSL) is 0.36.

Tronci*et al.*[6] has worked on a CBIR based tool named as Image hunter. The proposed tool uses the relevance feedback for the CBIR result retrieval. Also the authors have aimed at making it user friendly which enables almost all users to use it, not only the technical experts. In this paper the techniques of relevance feedback have been implemented which is based on k-NN and SVM relevance feedback. The “ImageHunter" tool evaluates the user feedback in order to optimize the image results and attained high performance by browsing the image database. The results calculated by using parameters precision and recall. However the SVM exhibits better performance as compared to k-NN.

Mussarat Yasmin *et al.* [16] has collectively developed a powerful image descriptor based upon shape, color and relevance feedback. The shape and color features have been used to fetch the matching images, whereas the relevance feedback is obtained from the users in order to optimize the final results produced by the duo algorithms in the latter portion of CBIR. Canny edge detector was used for detecting the shape of images and Color Histogram was used to calculate the probability mass function. The proposed image descriptor results in higher degree of relevancy and good balance between precision and recall.

Jing Yu *et al.*[14] has worked upon the bag-of-features model based CBIR. In this paper fusion of two algorithms designed, SIFT-LBP and HOG-LBP. The authors have used the feature integration analysis to achieve the goal they have aimed at. There were two types of feature integration methods performed i.e. Patch-based feature integration and Image-based feature integration. Then the k-means clustering applied to both the methods.. The finals results shown that image based feature integration method performs better as compared to patch based integration method.

RuiHy*et al*. [10] have conducted a survey to evaluate the performance of gradient field HOG descriptor. The authors have proposed the use of the gradient field HOG within a bag of visual words framework for the sketch based image retrieval. The image database has been collected from the Flicker of 15k images and these images were divided into 33 categories for the evaluation of sketch based image retrieval (SBIR).The proposed technique GF-HOG compared with other existing local descriptors i.e. SIFT, SSIM, HOG, Structure Tensor and Shape context. The experimental results show that GF-HOG descriptor outperforms other state of art descriptor on the Flicker dataset. The mean average precision score of GF-HOG is 7.35%.

Xiaoyu Wang *et al.*[11]suggested a novel approach of human detection for the purpose of handling partial occlusion. In this approach, HOG and LBP integrated to extract features. SVM act as a classifier for human detection. The experimental results performed on INRIA dataset. The HOG-LBP detector performs both with occlusion and without occlusion human detection. The evaluation based on two criteria False positive per window

(FPPW) and False Positive per image(FPPI).The purposed HOG-LBP detector achieves 20% better performance as compared to existing state of art descriptors such as Multilevel HOG feature , Shape and Haar wavelet feature etc.

III. THE APPROACH

The CBIR algorithms consist of following four phases: Dataset preparation, Preprocessing, Feature extraction and similarity matching. These phases performed in same manner except feature extraction phase in all algorithms. The generalized steps of any CBIR algorithm has been shown in Figure 1.

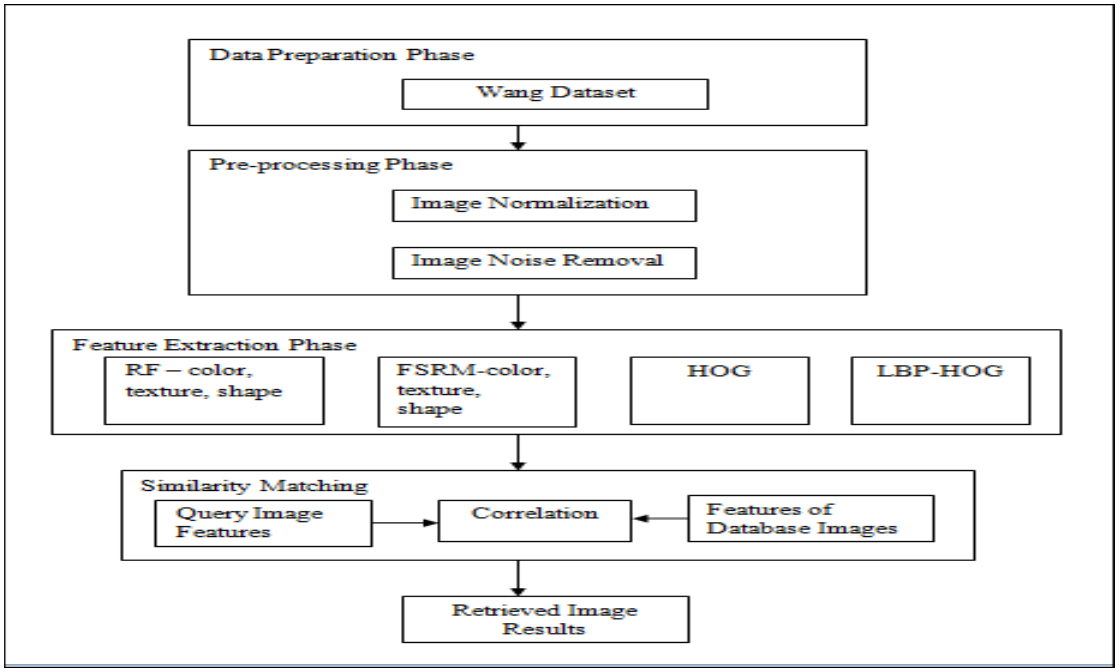

Figure 1: Generalized steps of CBIR Algorithm

A. Dataset Preparation

Wang dataset is a standard benchmark for content based image retrieval which is a subset of Corel stock image database [13]. It contains a total of 1000 images for 10 different categories like tribe, beach, buildings, buses, dinosaurs, elephants, roses, horses, mountains, food. The images have been taken under different illumination condition and different viewpoints. In this paper, 5 classes of Wang dataset for training and testing i.e. tribe, beach, buildings, buses and dinosaurs have been used. These 5 classes are used for relevance estimation, a query image is given and it is assumed that user is searching for images from the same class, and therefore the remaining 99 images from the same class are considered relevant and images from all other classes are irrelevant.

B. Preprocessing

In preprocessing, there are mainly two steps performed i.e. image normalization and image enhancement. At first, image is converted from RGB to grey scale. Each grey scale image is resized to 120*80. For image normalization, Gaussian filter is used to remove noise from image and helps in smoothens the image. Histogram equalization is used for image enhancement to improve the quality of the images.

C. Feature Extraction

The presented study of content based image retrieval consists of four algorithms i.e. CBIR-Relevance Feedback, CBIR- fuzzy semantic relevance matrix, CBIR-HOG, CBIR- LBP HOG. The first two techniques which are based on relevance feedback uses two parameters, color and texture for feature extraction [14], as these parameters are important to retrieve the similar images from database. The other two techniques are CBIR-HOG and CBIR-LBP HOG which uses HOG[15] and LBP-HOG[16] for feature extraction respectively.

Color features are extracted by using Color histograms. Color histogram used as local descriptor to calculate the color intensity of particular image. RGB and HSI color models are used for assigning color values. To extract the texture features, entropy, local range and standard deviation measures are used as parameters. After extraction of these features, a feature vector is created consists of color and texture feature values.

The extraction of HOG features occurs in various steps: The first step is to calculate the gradient values by applying derivative mask in both horizontal and vertical directions. The computation of gradient values calculates in form of magnitude $|G|$ and orientation θ.

$$|G|=\sqrt{I^2x + I^2y} \qquad (1)$$

$$\theta = \arctan\frac{I_y}{I_x} \qquad (2)$$

Where I_x and I_y are derivatives of an image I.

After computation of gradients, the second step is orientation binning. In this step, cell histograms created by calculating pixels of an image. Each pixel within cell has weighted value which found in computation of gradient. The HOG descriptor is then vector of components of the normalized cell histograms from all of the block regions. To normalize the orientation histograms, they should be grouped into blocks which are called as block normalization. Thus the HOG feature is extracted from an image.

In LBP-HOG, The value of LBP is computed by calculating the gray level value of the center pixel and its neighbors in the give image. The following equation is used for computation of LBP descriptor.

$$LBP_{P,R} = \text{p=1P2 (p-1) f1 } (g_{p-}g_c) \qquad (3)$$

$$\text{f1(x) = 1, x}\geq 00\text{, otherwise}$$

where P is no of neighbors and R is center of the neighborhood. $g_c$is gray value of center pixel and g_p is gray value of its neighbors.

When the LBP features are extracted, combine both HOG and LBP features and create a feature vector.

D. Similarity Matching

The similarity matching as name suggests are parameters to determine the similarities between different feature vectors [10]. The information gained by feature extraction is used to measure the similarity between two images. The similarity measure Correlation Coefficient is used [18]. After the extraction of feature vectors, Correlation

Coefficient between two feature sets is calculated. If the distance between the two feature set is minimal then images are more similar and vice versa.

Equation to calculate Correlation corr2 is defined by:

$$corr2 = \frac{\sum_m \sum_n (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{(\sum_m \sum_n (A_{mn} - \bar{A})^2)(\sum_m \sum_n (B_{mn} - \bar{B})^2}} \quad (4)$$

Where A_{mn} and B_{mn} are feature vectors

$\bar{A}$ = mean (A) and $\bar{B}$ = mean (B)

IV. ALGORITHM

i. CBIR- Relevance feedback

In relevance feedback, color and texture features are used for feature extraction and similarity matching between query image and database image is calculated by using Correlation coefficient metric. The retrieved images displayed on screen. The user decides whether the results are relevant or not. If the results found to be irrelevant, feedback loop is repeated many times until the user satisfied.

Algorithm : Relevance Feeback	Algorithm: CBIR FSRM
1. *Read RGB image I_m.* 2. *Convert I_m to grey scale.* 3. *Apply Gaussian filter G_f on I_m.* 4. *Extract color F_c & texture F_t features of I_m* 5. *Create feature vector F_{ct}* 6. *Compute Correlation Coefficient C_f for similarity matching.* 7. *If results are irrelevant,* *Repeat step 4 to 7* *Else* *Stop*	1. *Read RGB imageI_m.* 2. *Convert I_m to grey scale.* 3. *Apply Gaussian filter G_f on I_m.* 4. *Extract color F_c & texture F_t features of I_m* 5. *Create feature vector F_{ct}* 6. *Compute Correlation Coefficient C_f for similarity matching.* 7. *Build the FSRM and display results* 8. *If user not satisfy* *Set the irrelevant image to bit 0.* *Repeat step 4 to 8* *Else* *Stop.*

ii. CBIR- FSRM

Fuzzy semantic relevance feedback is constructed to describe semantically relevance between the images in the database. The weights in the FSRM are adjusted according to user's feedback. A fuzzy set is a class of objects with a continuum of grades of membership. The value of fuzzy set lies in the interval [0, 1]. Therefore, we use the value of interval [0, 1] represents the "grade of membership" of an object of the concept. The more the value closes to 1, the more the image similar to query image. The working of FSRM explains as follows:-

iii. CBIR- HOG

In this approach, Histogram of Oriented Gradients features are computed by calculating gradient of color of images. The histograms of each gradient image are created, and then normalize these histograms using block normalization

method. Correlation coefficient is calculated between query image and testing images. Thus the retrieval results are displayed on screen. The following is the algorithm for HOG algorithm.

Algorithm: CBIR-HOG	Algorithm: CBIR-LBP HOG
1. *Read RGB image* I_m. 2. *Convert* I_m *to grey scale.* 3. *Apply Gaussian filter* G_f *on* I_m. 4. *Calculate gradient* G_c *of* I_m 5. *Create Histogram* H_e *of each* G_c 6. *Apply Block normalization* B_e *on* H_e 7. *Create feature vector* F_h 8. *Compute Correlation coefficient* C_f 9. *Results displayed.* 10. *End*	1. *Read RGB image* I_m. 2. *Convert* I_m *to grey scale.* 3. *Apply Gaussian filter* G_f *on* I_m. 4. *Compute LBP descriptor* L_d *of* I_m. 5. Create feature vector F_l. 6. Merge values of F_h *and* F_l. 7. *Compute Correlation Coefficient* C_f 8. *Results displayed* 9. *End*

iv. CBIR- LBP HOG

In this approach, Local Binary pattern (LBP) features are concatenated with HOG features for content based image retrieval. At first, LBP feature descriptor is computed for each image, then created a feature vector of LBP features. This feature vector combines with already derived HOG feature vector. Correlation coefficient is calculated between query image and testing images. Thus the final results displayed on screen. The following is the algorithm for LBP-HOG approach.

V. EXPERIMENTAL RESULTS

In this section, the detailed analysis of four algorithms is provided. The standard WANG image database is used for experimental results. The MATLAB environment has been used to implement algorithms. The Support Vector Machine (SVM) is used for training and classification[19]. The performance of the system is evaluated using performance parameters i.e. sensitivity, specificity, negative predictive value (NPV), precision, accuracy, overall accuracy and Kappa analysis [20]. These parameters are calculated by using True positive (TP), True negative (TN), False positive (FP), False negative (FN) values.

The results have been obtained for each algorithm. Among the four algorithms CBIR-HOG depicts the better performance in terms of all parameters as compared to other three algorithms. The performance of HOG feature descriptor is extremely good in CBIR mechanism. Because of the window based descriptor it extracts features more precisely and results comes out to be more relevant as compare to other algorithms. Local Binary Pattern feature descriptor when integrated with HOG performs better only in building class than individual HOG descriptor. The Relevance Feedback and FSRM also exhibit good performance but not as well as HOG or LBP-HOG. The average values of precision, sensitivity, specificity, NPV, accuracy and overall accuracy have been calculated to evaluate the overall performance of the system. According to Kappa analysis all the four algorithms attains substantial performance. The table 1 summarizes the performance in terms of Precision, Sensitivity, Specificity, Accuracy, Overall accuracy and Kappa

Table 1: Overall performance of CBIR algorithms

Algorithm	Sensitivity	Specificity	Precision	Negative Predictive value	Accuracy	Overall Accuracy	Kappa Analysis
FSRM	80.67%	93.18%	78.24%	94.65%	92.48%	78.84%	0.735
RF	78.72%	96.34%	77.81%	94.13%	91.7%	78.31%	0.732
HOG	84.56%	96.47%	86.13%	96%	93%	83.52%	0.779
LBP-HOG	81.46%	95.8%	85.6%	95.8%	92.75%	81.29%	0.751

The average precision, specificity, sensitivity, NPV and overall accuracy of each algorithm have been illustrated in Figure 2 with the help of line graph. The average precision is highest for HOG and while RF holds the lowest value. The sensitivity or recall is also highest in case of HOG. There is very little variation in case of sensitivity and NPV, but HOG attains highest value for sensitivity. The overall accuracy is highest for HOG and LBP-HOG attains less comparatively HOG whereas FSRM and RF hold almost same value with little difference.

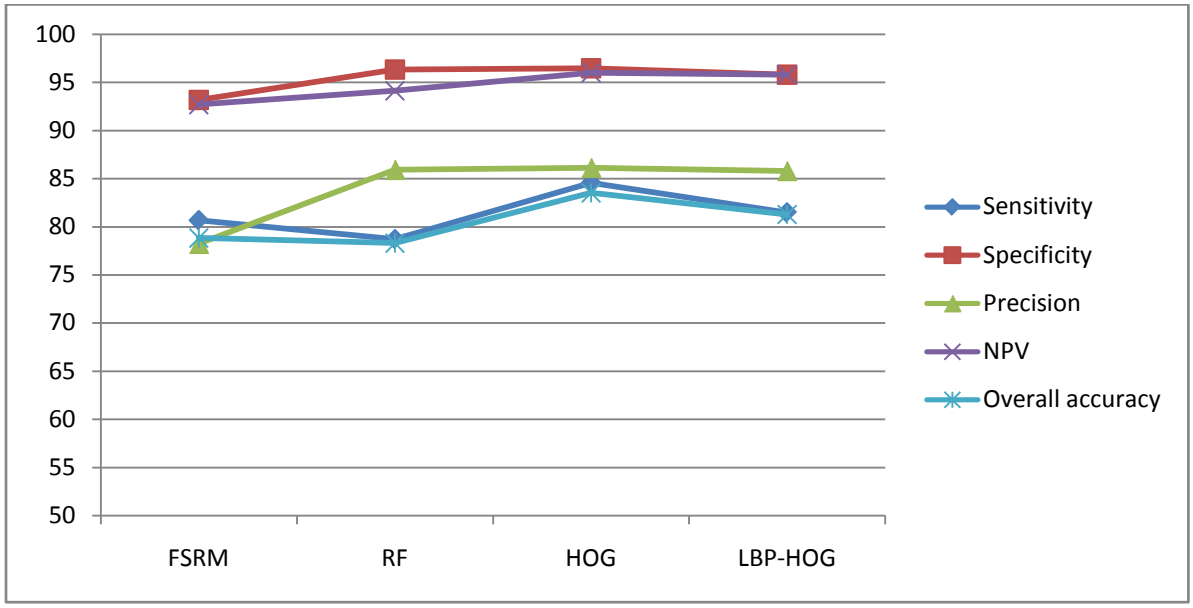

Figure 2: Performance of CBIR Algorithms

VI. CONCLUSION

Content based image retrieval is a dominant way of retrieving images. In recent years, many researchers have proposed various techniques for the content based image retrieval. A considerable number of different feature extraction techniques and similarity measures have also been proposed.

The presented algorithms for CBIR have varying impact on the results of image retrieval. These algorithms for content based image retrieval were evaluated for five classes of images i.e. tribal, beach, building, bus and dinosaur. The results fortified that the HOG and LBP-HOG feature extraction algorithm is more efficient as compared to Relevance Feedback and FSRM algorithms.

REFERENCES

[1] D. S. Shete and M. S. Chavan, "Content Based Image Retrieval : Review," *Int. J. Emerg. Technol. Adv. Eng.*, vol. 2, no. 9, pp. 85–90, 2012.

[2] H. Narasimhan and P. Ramraj, "Contribution-based clustering algorithm for content-based image retrieval," *in Proceedings of International Conference on Industrial and Information Systems*, pp. 442–447, 2010

[3] Y. Gao, M. Wang, Z. J. Zha, J. Shen, X. Li, and X. Wu, "Visual-textual joint relevance learning for tag-based social image search," *IEEE Transactions on Image Processing*, vol. 22, no. 1, pp. 363–376, 2013.

[4] H. Aboulmagd, N. El-Gayar, and H. Onsi, "A new approach in content-based image retrieval using fuzzy," *Telecommunication Systems*, vol. 40, no. 1–2, pp. 55–66, 2008.

[5] R. Tronci, G. Murgia, M. Pili, L. Piras and G. Giacinto, "ImageHunter: A Novel Tool for Relevance Feedback in Content Based Image Retrieval," In: *Lai C., Semeraro G., Vargiu E. (eds) New Challenges in Distributed Information Filtering and Retrieval. Studies in Computational Intelligence, vol 439. Springer, Berlin, Heidelberg*, pp 53-70, 2013

[6] L. Pinjarkar, M. Sharma, and K. Mehta, "Comparison and Analysis of Content Based Image Retrieval Systems Based On Relevance Feedback," *Journal of Emerging Trends in Computing and Information Sciences*, vol. 3, no. 6, pp. 833–837, 2012.

[7] L. Wang and W. Lin, "A Semantic Subspace Learning Method to Exploit Relevance Feedback Log Data for Image Retrieval," *IEEE Symposium on Computational Intelligence and Data Mining (CIDM),* pp. 178–183., 2013

[8] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR"05) - Volume 1*, pp. 886–893, 2005.

[9] Y. Sheng, L. Xianmei, and UK Borsay, "A Pedestrian Detection Method Based on the HOG-LBP Feature and Gentle AdaBoost," *International Journal of Advancements in Computing Technologies.*, vol. 4, no. 19, pp. 553–560, 2012.

[10] R. Hu and J. Collornosse, "A performance evaluation of gradient field HOG descriptor for sketch based image retrieval," *Journal of Computer Vision and Image Understanding*. vol. 117, no. 7, pp. 790–806, 2013.

[11] X. Wang, T. X. Han, and S. Yan, "An HOG-LBP human detector with partial occlusion handling," in *Proceedings of IEEE International Conference on Computer Vision*, pp. 32-39, 2009.

[12] J. Zhang, K. Huang, Y. Yu, and T. Tan, "Boosted local structured HOG-LBP for object localization," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1393–1400, 2011.

[13] J. Li and James Z. Wang, "Wang dataset", *http://wang.ist.psu.edu/docs/related*

[14] J. Yu, Z. Qin, T. Wan, and X. Zhang, "Feature integration analysis of bag-of-features model for image retrieval," *Neurocomputing*, vol. 120, pp. 355–364, 2013

[15] X. Yuan, J. Yu, Z. Qin, and T. Wan, "A SIFT-LBP image retrieval model based on bag of features," in *Proceedings of International Conference on Image Processing*, pp. 1061–1064, 2011.

[16] M. S. R. Janani and P. Sebhakumar, "An Improved CBIR Method Using Color and Texture Properties with Relevance Feedback," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 2, no. 1, pp. 47–54, 2014.

[17] D. C. G. Pedronette and R. Da S. Torres, "Distances correlation for re-ranking in content-based image retrieval," in *Proceedings of SIBGRAPI Conference on Graphics, Patterns and Images,* pp. 1–8, 2010.

[18] A. Marakakis, N. Galatsanos, A. Likas, and A. Stafylopatis, "Relevance Feedback for Content-Based Image Retrieval Using Support Vector Machines and Feature Selection, ". *In: Alippi C., Polycarpou M., Panayiotou C., Ellinas G. (eds) Artificial Neural Networks – ICANN 2009. ICANN 2009. Lecture Notes in Computer Science, vol 5768. Springer, Berlin, Heidelberg*, pp. 942–951, 2009.

[19] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *Biometrics*, vol. 33, no. 1, pp. 159–174, 1977.