Content-Based Image Retrieval using Local Patterns and Supervised Machine Learning Techniques

Maher Alrahhal[1], Supreethi K.P.[2]

[1,2]CSE Dept., JNTUH Collage of Engineering, Hyderabad 500085, Telangana, India
[1]maherrahal92@gmail.com, [2]supreethi.pujari@gmail.com

Abstract: CBIR is a very important domain, especially in the last decade due to the increased need for image retrieval from the multimedia database. In general, we extract low level (color, texture, and shape) or high-level features (when we include machine learning techniques) from the images. In our work, we proposed a new CBIR system using Local Neighbor Pattern (LNP) with supervised machine learning techniques. We evaluated the performance of this system by comparing the system with Local Tetra Pattern (LTrP) using Corel 1k, Vistex and TDF face databases. We used three types of the database (i.e color, texture and face databases) to improve the effectiveness of our system. Performance analysis shows that LNP gives better performance regarding the average recall than LBP, LDP, and LTrP. To increase the accuracy of this system we used the LNP method with machine learning techniques and performance analysis shows that local pattern with machine learning techniques improves the average accuracy from 36.23% to 85.60% when we use LNP with cubic SVM on DB1 (Corel1K), and from 82.51% to 99.50 % when we use LNP with fine KNN on DB2 (Vistex DB), and from 56.63% to 95% when we use LNP with ensemble subspace discriminant on DB3 (face DB).

Keywords: Content-Based Image Retrieval, Local Neighbor Patterns, Supervised Machine Learning Techniques, Classification, MATLAB.

I. INTRODUCTION

Image Retrieval (IR) means the process of searching the related images in images Database, and retrieve the most similarity image to the user. IR techniques can be classified into Text-Based or Content-Based Image Retrieval. TBIR is the process of manually adding annotation or description to the image in the database to describe the content of images and sometimes for describing other metadata of images like size, dimensions, and format of images. The advantages of the TBIR system are simple and very fast in displaying the result and depend on matching a textual query with description of the images. However, this method has many disadvantages, such as the error rate is high and a great amount of manpower and material resources are needed [1]. Also, adding a description of the image depend on our points of view, or how we understand the image, so the description of the image may differ from one person to another.

Finally, the big issue is adding an annotation to the image depending on your languages.

CBIR is a method that is used to find an image from a set of the large database as per the user requirement. CBIR is also known as query by image content (QBIC) [2]. CBIR ([3], [4], [5], [6], [7]) is an alternative to the traditional TBIR which overcomes the above limitations. CBIR covers several areas like image segmentation, extracts a feature from the image and converts this feature into a semantic feature. In general, we extract low level (color, texture, and shape) or high-level features (when we include machine learning techniques) from the images [8]. In general, CBIR consists of two phases [9] off-line phase for feature extraction and online phase for image retrieval. In the offline phase the system extracts the feature from all the images in the DB and store them in DB. In on-line phase, the user inputs a query and the system extracts the features from this image and measure the similarity by calculate the distance between query image (i.e feature vector) and all images in DB (i.e feature vectors), then sort the distance in ascending way and retrieve the top k images to the user. Another technique in image retrieval is combining text-based and content- based to increase accuracy and get better performance.

CBIR has been used in several fields, such as satellite images [10], remote sensing, medical imaging[11], fingerprints scanning ([12], [13]) and biodiversity information systems. CBIR techniques are being used in the area of satellite images to find earth minerals, aerial survey, for monitoring agriculture, to generate weather reports and for tracking surface objects. Medical imaging is one of the prominent areas of application of CBIR which can be used for monitoring patient health reports, to aid diagnosis by identifying the similar past cases etc. When given a fingerprint query image CBIR systems can be used to extract the similar fingerprint images that results in verification of an individual. Fingerprints are used in banking sector,colleges, corporate companies, and forensic labs .

II. RELATED WORK

For many years, researchers have been using different local descriptor methods for image retrieval and classification. Image classification is supervised machine learning, for that in training step, we have data sample with a label (class name) for

978-1-5386-9346-9/19/$31.00 ©2019 IEEE

training the classifier. In the test stage, we have data sample as input to the classifier and the classifier will give us class name as output. For that, image classification represents as black box function which takes input (animals as for example) and returns the text label of animals (cat as an example). And training step here is finding best function to achieve this operation. Currently, retrieval systems tend to extract features from images using (e.g. LBP[14], LDP[15] and LTrP[16]) then train machine learning classifiers to classify images into predefined semantic classes, finally test the model by classifying the new image. The most important point is to evaluate classification results which are often better than image retrieval using feature extraction only. Priyanka Pawar et al proposed Local Binary Pattern (LBP) [14]. LBP is a very good method for texture images. This method works as follows, first, we convert the color image into a grayscale image, then compare center pixel of 3 x 3 window with each neighbor, If the value of neighbor equal or greater than a center pixel, then we replace the neighboring pixel value by 1, otherwise, replace it by 0. Finally, convert those binary value to decimal value and replace center value by decimal value. For getting the feature vector, we calculate the histogram of the new decimal matrix with 256 bins. The variants of LBP, such as global matching [17], dominant LBPs [18], completed LBPs [19] are proposed for rotational invariant texture classification. LBP is local pattern which consider as non-directional (because encodes all-direction)first-order derivative operator, while LDP [15],[20] calculate higher-order(second, third or more) derivative information along 0°,45°,90°, and 135° and the feature vector size for each direction is 59, so size of feature vector of LDP is 4×59 =236 which contains more detailed and give better result than LBP. S. Murala et al proposed Local Tetra Patterns (LTrP) [16] which encodes the relationship between the center pixel and its neighbors, based on the directions and we calculate the directions using the first-order derivatives in both directions (vertical and horizontal). Md Mahmudur et al, [21] proposed the probabilistic outputs of multi-class support vector machine classifier (SVM) to predict query image and use image information to remove irrelevant images and re-set the feature weights in similarity matching. Mina et al, [22] proposed a medical decision support system, which extract the features from the medical images using GLCM (gray level co-occurrence matrices), then use PCA (principal component analysis) to reduce the size of feature vector and classify the images using SVM (support vector machine) to discover images with tumor and image of multiple sclerosis.

III. SUPERVISED MACHINE LEARNING TECHNIQUES

In supervised machine learning, we divide the dataset into two parts, first part for the train the system and the rest of images for testing the system. After the training step completed, the system must have the ability to give accurate results when given new data for testing. In this work, we use four types of supervised machine learning algorithms to classify images.

A. Linear Discriminant

Linear discriminant [23] separates between classes using linear boundaries by estimates the parameters of Gaussian distribution (i.e mean and variance) where linear discriminant assume that data subject to Gaussian distribution. The linear discriminant is a fast algorithm for classification and good for wide datasets.

B. Support Vector Machine

Classifies data by finding the best hyperplane [24] (has a maximum margin) which separates one class from other classes. Margin means the maximum gaps between classes without containing any data points, and support vectors mean data points which are nearest to hyperplane as shown in fig.1. Many types of SVM classifiers are used depends on kernel functions: 1-Linear kernel: classify points by finding the linear separation between classes. 2- Medium Gaussian SVM: which is characterized by medium distinctions, with kernel scale set to *sqrt (p), 3*- Quadratic kernel and 4- Cubic kernel.

C. K-Nearest Neighbor classification

Given a set of points and distance function, K-nearest neighbor (KNN) [25] finds the K closest points to a query point or set of points. Many types of KNN classifiers differ from each other in different distance between classes and number of neighborhood, are used:

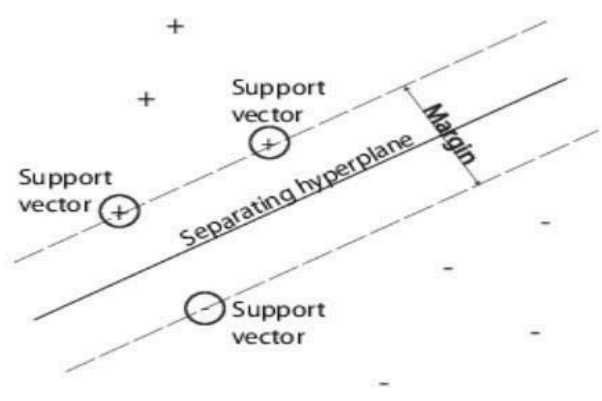

Fig. 1. Example of SVM

1- Fine KNN: finely detailed distinctions between classes, where K= 1, 2- Medium KNN, use medium distinctions between classes, with K= 10. Cosine KNN, Cubic KNN, and Weighted KNN use medium distinctions between classes, with cosine, cubic and distance weight metric, and K(No. of neighbors)= 10.

D. Ensemble Classifiers

Combine many weak learners to make high-quality ensemble model [26]. Qualities depend on the choice of algorithm. Many types of ensemble classifiers are used: 1- Ensemble bagged trees: we use random forest bag, with decision tree learners, 2- Ensemble subspace discriminant: we use subspace, with discriminant learners, 3- Ensemble subspace KNN: we use subspace, with nearest neighbor learners.

In the classification step, we consider k fold cross validation that means we divided the data into k partitions, k-1 parts for training and one for validation (testing the model). This process repeated k times that means each of the k parts used as training and testing. Finally, we take the average accuracy of k times to get the final accuracy of the model. Usually, k consider 10 [27] and in my work we take k=10.

IV. METHODOLOGY

A. *Feature Extraction using Local Neighbor Method*

This method depends on converting the image into grayscale image and apply 3×3 window on the image, then change center pixel value in a 3×3 window depending on neighbors of the center pixel and repeat this procedure to include the entire picture. Finally, construct a feature vector by taking the histogram with 256 bins. We change center pixel value of 3×3 window as follows and shown in fig.2:

$V_1=B_2+B_8+B_0-3B_1$

$V_i=B_{i+1}+B_{i-1}+B_0-3B_i$ i=2...7

$V_8=B_1+B_7+B_0-3B_8$

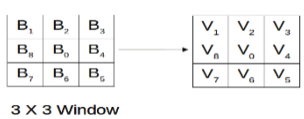

Fig. 2 (3×3) Window and equations to encode the center pixel

Check V_i values, if $V_i<0$, replace it with 0; otherwise replace it with 1, finally convert V_i values to a decimal value by considering V_1 as LSB and V_8 as MSB; i=1...8. Fig.3 *illustrates an example of LNP method.*

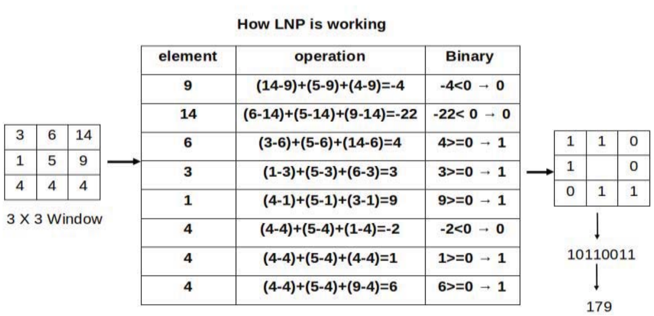

Fig. 3 Example of LNP method

B. Proposed CBIR with machine learning system

This system consists of two phases as shown in fig.4: offline phase for extract feature vector for all images in the database then get feature vectors for 60-70 % images from each class and add a label (class name) for those features then train machine learning classifier. Finally, predict class name for all feature vectors via machine learning classifier.

In online phase, the user enters the query and the system extracts LNP features and predicts to which class it belongs to using machine learning classifier, finally the system retrieves the images which classified as same query image's class in offline phase to the user.

1. Off-Line Phase

a) For each image in the database calculate feature vector using Local Patterns methods.

b) Get feature vectors for (60-70) % images of each class in the database.

c) Add a label (class name) for each feature vectors.

d) Train machine learning classifier.

e) Predict class name for each feature vector in the database and get class name for each image in the database.

Note:

Machine learning classifiers mean Linear Discriminant or different techniques of SVMs or different techniques of KNNs or different techniques of Ensemble methods.

For feature extraction, we used the LNP method.

2. On-Line Phase

a) Read the query image from the user interface.

b) Calculate the feature vector for the query using LNP methods.

c) Predict class name for query image using machine learning classifier which learned in offline phase, says "S1".

d) Get feature vectors for all images which classified in offline phase as “S1”.

e) Calculate the Euclidean distance between the query feature vector and all feature vectors in the last step.

f) Sort this distance in ascending way.

g) Retrieve top K images, where the user chooses the value of K.

V. EXPERMINTAL RESULTS AND DISSCUSSIONS

We used three different types of databases for experimental evaluation: Color database (Corel 1k) [28], Texture database (Vistex database) [29] and faces database (TDF) [30]. In the first Experiment, We compared our proposed method LNP with LBP, LDP, and LTrP in term of average recall. In the second experiment, we compare between CBIR system using LNP VS. CBIR system using LNP and machine learning techniques. To evaluate our proposed method, we implement and test the effectiveness of this system in MATLAB R2017b.

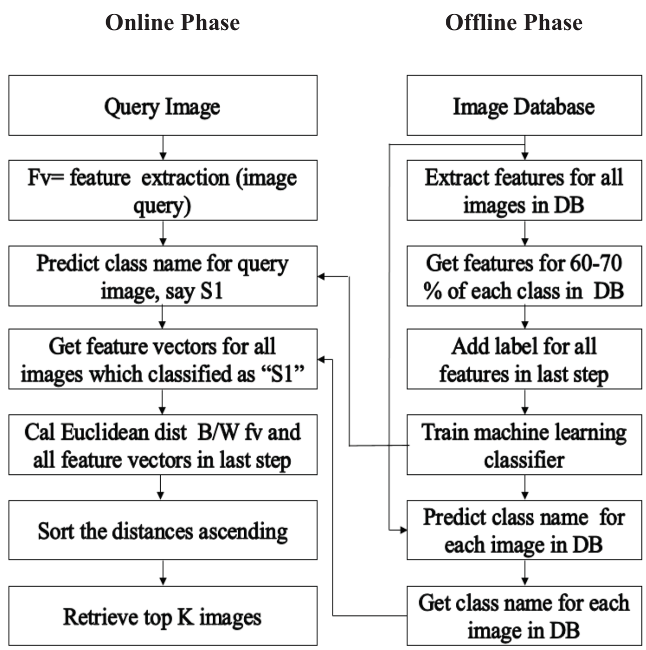

Fig. 4. Flowchart of the proposed method CBIR method with ML Techniques

A. Datasets

1. Color Database (Corel 1K)

Corel 1k database [28] consists of thousand color images of size 384×256 or 256×384. Corel 1k consists of 10 classes. These classes are shown in fig.5 and each class consists of 100 images.

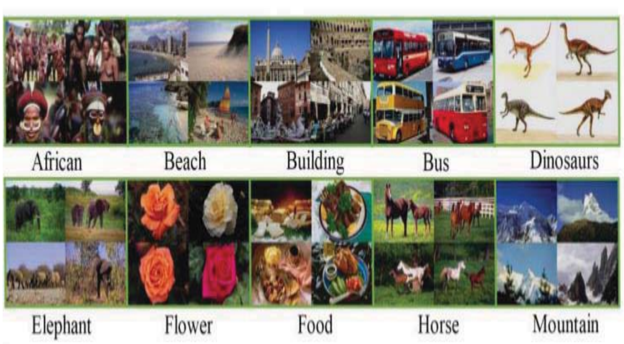

Fig. 5. Corel 1K classes

2. Vistex Database

Vistex database [29] consists from 40 images. The size of each image is 512×512 which is divided into 16 non-overlapping images of size 128×128. This database consists of 40×16=640. Fig.6 illustrates Vistex database classes.

Fig. 6. Vistex database classes

3. Faces Database

The database of faces (TDF) [30] consists of 40 classes. Each class belong to one person and each class contains 10 images and the size of each image is 112×92 pixels. Those images differ in lighting, facial details and facial expressions . Fig.7 illustrates one image from each face database classes.

Fig. 7. One image from each face database classes

B. Performance Measures

In experiment 1, we compared between local pattern methods in term of average recall performance measure [31] which calculate as:

Recall=Number of Images related (classified) to query's class divided by a total number of images in query's class.

For calculate Recall, the user enters query image to CBIR system which extracts the features from query image using one of the local pattern methods (LBP, LDP, LTrP or LNP), then measure the distance between this feature query and all features in the DB, then CBIR system retrieve top K images to the user. In this paper, we consider K= max No. of images in the class of DB (i.e K=100 for Corel 1k database, K=16 for Vistex DB and K=10 for face database). We repeat this procedure for all images in the database and then we calculate the average, to get an average recall. The best recall is 1.0 or (100% as a percent), whereas the worst is 0.0.

In experiment 2, we consider the accuracy as performance measure which is calculated as:

Accuracy= No. of all images classified correctly ÷ the total No. of images in the class.

To calculate the accuracy, the user enters a query image to CBIR system and extract the features using local pattern method (LNP) then we classified the feature vector of query to which class it belongs to (say, image class is "S1") using machine learning classifier which is trained in off-line phase. Finally, we retrieve all images which classified as (S1) in the off-line phase to the user. The best accuracy is 1.0 or (100% as a percent), whereas the worst is 0.0. *If we look to average recall and average accuracy definition, we can observe that both are same.*

C. Results

Experiment 1

In this experiment , images from the three databases DB1,DB2 and DB3 are used and the performance is measured for the proposed local patterns in term of recall with Euclidean distance. In table 1, it is evident that LNR yield better results comparing with LBP, LDP, LTrP regarding of average recall. Where the feature vectors length are shown in table 2.

TABLE 1: Average Recall of Local Pattern

Experiment 1	Average Recall		
	Corel 1K DB	Vistex DB	face DB
LBP	33.96%	80.96%	54.28%
LDP	35.19%	77.27%	45.78%
LTrP	34.56%	78.79%	48.35%
LNP	**36.23%**	**82.51%**	**56.63%**

TABLE 2: Feature Vector Length

local pattern method	length
LBP	256
LDP	236
LTrP	767
LNP	256

Experiment 2

In experiment 2, images from DB1,DB2 and DB3 are used and the performance is measured for the proposed LNP in term of average accuracy. In table 3 and figures 8, it shows that the IR system with LNP plus machine learning techniques yields better performance comparing with image retrieval with LNP only. Performance analysis shows that LNP with machine learning techniques improves the average accuracy from 36.23% to 85.60% when we use LNP with cubic SVM on DB1 (Corel1K), and from 82.51% to 99.50 %when we use LNP with fine KNN (or LNP with linear discriminant) on DB2 (Vistex DB), and from 56.63% to 95% when we use LNP with ensemble subspace discriminant on DB3 (face DB).

TABLE 3: Comparing Between CBIR (LNP) Accuracy VS. CBIR + ML Techniques Accuracy

Databases		Corel 1K DB	Vistex DB	face DB
CBIR with LNP only		36.23%	82.51%	56.63%
Linear Discriminant		80.10%	**99.50%**	80.00%
LNP +KNN	Fine	71.20%	**99.50%**	91.80%
	Medium	72.60%	98.40%	77.20%
	Cosine	70.40%	97.20%	68.50%
	Cubic	71.70%	97.30%	78.80%
	Weighted	75.40%	98.90%	84.50%
LNP+SVM	Linear	82.30%	98.40%	86.00%
	Quadratic	85.40%	98.80%	86.00%
	Cubic	**85.60%**	98.40%	84.20%
	medium Gaussian	81.40%	99.20%	91.20%
Ensemble + LNP	Bagged Tree	74.10%	98.10%	69.80%
	Subspace Discriminant	83.10%	99.40%	**95.00%**
	Subspace KNN	74.40%	99.20%	86.50%

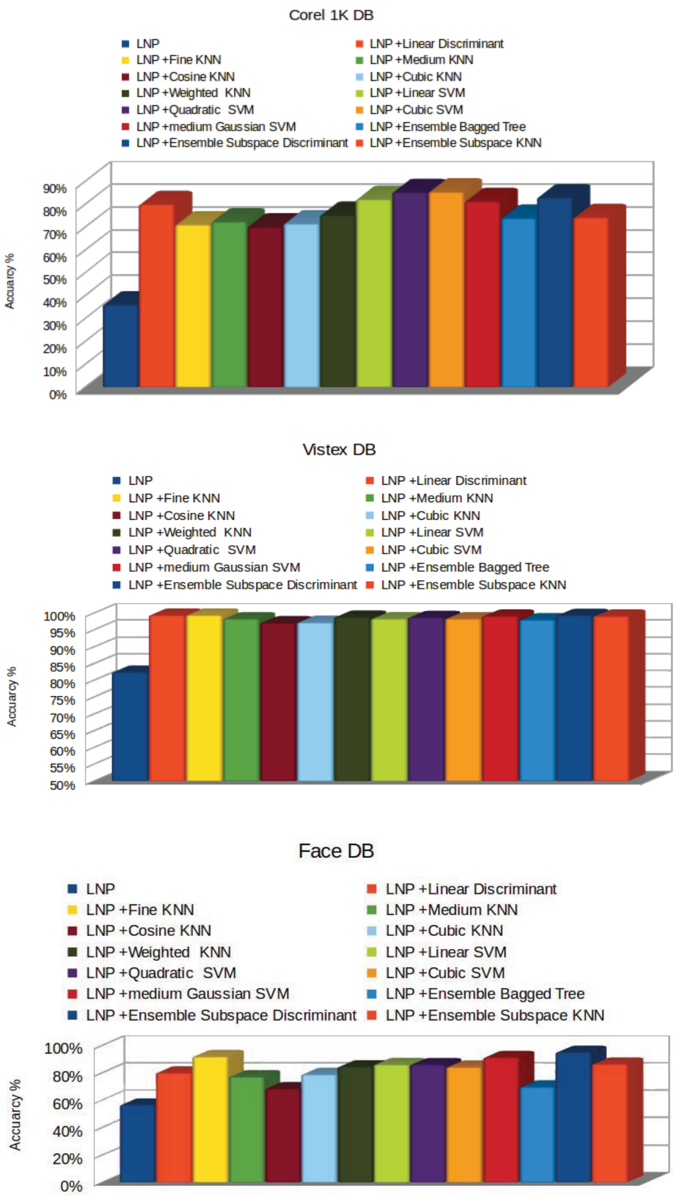

Fig. 8. Comparing between CBIR accuracy VS. CBIR + ML techniques accuracy on 1. Corel 1k, 2. Vistex and 3. Face databases

VI. CONCLUSION AND FUTURE SCOPE

In the proposed work we focused on two areas. Firstly, we proposed the LNP method for image retrieval which yields a better result when compared with LBP, LDP and LTrP methods in terms of average recall. Next, we focused on CBIR with machine learning techniques to improve the performance of the system. In the future, we can extend the work to develop the LNP method in video retrieval. Deep learning techniques can be used in CBIR to acquire more accuracy and minimize training time for the machine. Also, we can combine CBIR with Hadoop techniques to process huge images database and working in a distributed environment.

REFERENCES

[1] T. He, Y. Wei, Z. Liu, G. Qing and D. Zhang, "Content Based Image Retrieval Method Based on SIFT Feature", in *Proc. ICITBS*, Xiamen China, pp. 649-652, 2018.

[2] S. Sharma, "Use of Artificial Intelligence Algorithm for Content-Based Image Retrieval System", *International Journal of Advance Research, Ideas and Innovations in Technology*, vol. 4, no., pp. 680-684, 2018.

[3] Y. Rui, T. Huang and S. Chang, "Image Retrieval: Current Techniques, Promising Directions, and Open Issues", *Journal of Visual Communication and Image Representation*, vol. 10, no. 1, pp. 39-62, 1999.

[4] A. Smeulders, M. Worring, S. Santini, A. Gupta and R. Jain, "Content-based image retrieval at the end of the early years", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 12, pp. 1349-1380, 2000.

[5] M. Kokare, B. Chatterji and P. Biswas, "A Survey on Current Content based Image Retrieval Methods", *IETE Journal of Research*, vol. 48, no. 3-4, pp. 261-271, 2002.

[6] Y. Liu, D. Zhang, G. Lu and W. Ma, "A survey of content-based image retrieval with high-level semantics", *Pattern Recognition*, vol. 40, no. 1, pp. 262-282, 2007.

[7] R. Datta, D. Joshi, J. Li and J. Wang, "Image retrieval", *ACM Computing Surveys*, vol. 40, no. 2, pp. 1-60, 2008.

[8] A. Nazir, R. Ashraf, T. Hamdani and N. Ali, "Content Based Image Retrieval System by using HSV Color Histogram, Discrete Wavelet Transform and Edge Histogram Descriptor", in *Proc. iCoMET*, Sukkur Pakistan, 2018.

[9] M. Alkhawlani, M. Elmogy and H. El Bakry, "Text-based, Content-based, and Semantic-based Image Retrievals: A Survey", *International Journal of Computer and Information Technology (ISSN: 2279 – 0764)*, vol. 04, no. 01, pp. 58-66, 2015.

[10] A. Ferran, S. Bernabe, P. Rodriguez and A. Plaza, "A Web-Based System for Classification of Remote Sensing Data", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 4, pp. 1934-1948, 2013.

[11] J. Ramos, T. Kockelkorn, I. Ramos, R. Ramos, J. Grutters, M. Viergever, B. van Ginneken and A. Campilho, "Content-Based Image Retrieval by Metric Learning From Radiology Reports: Application to Interstitial Lung Diseases", *IEEE Journal of Biomedical and Health Informatics*, vol. 20, no. 1, pp. 281-292, 2016.

[12] J. Montoya Zegarra, N. Leite and R. da Silva Torres, "Wavelet-based fingerprint image retrieval", *Journal of Computational and Applied Mathematics*, vol. 227, no. 2, pp. 294-307, 2009.

[13] M. Gavrielides, E. Sikudova and I. Pitas, "Color-based descriptors for image fingerprinting", *IEEE Transactions on Multimedia*, vol. 8, no. 4, pp. 740-748, 2006.

[14] P. Pawar and P. Belagali, "Image Retrieval Technique Using Local Binary Pattern (LBP)", *International Journal of Science and Research (IJSR)*, vol. 4, no. 7, pp. 1440-1443, 2015.

[15] Baochang Zhang, Yongsheng Gao, Sanqiang Zhao and Jianzhuang Liu, "Local Derivative Pattern Versus Local Binary Pattern: Face Recognition With High-Order Local Pattern Descriptor", *IEEE Transactions on Image Processing*, vol. 19, no. 2, pp. 533-544, 2010.

[16] S. Murala, R. Maheshwari and R. Balasubramanian, "Local Tetra Patterns: A New Feature Descriptor for Content-Based

Image Retrieval", *IEEE Transactions on Image Processing*, vol. 21, no. 5, pp. 2874-2886, 2012.
[17] Z. Guo, L. Zhang and D. Zhang, "Rotation invariant texture classification using LBP variance (LBPV) with global matching", *Pattern Recognition*, vol. 43, no. 3, pp. 706-719, 2010.
[18] S. Liao, M. Law and A. Chung, "Dominant Local Binary Patterns for Texture Classification", *IEEE Transactions on Image Processing*, vol. 18, no. 5, pp. 1107-1118, 2009.
[19] Zhenhua Guo, Lei Zhang and D. Zhang, "A Completed Modeling of Local Binary Pattern Operator for Texture Classification", *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1657-1663, 2010.
[20] P. Reddy and K. Satya Prasad, "Local Derivative Patterns and their Magnitudes for Content Based Image Retrieval", *International Journal of Computer Science & Informatics*, vol. 1, no. 2, pp. 62-67, 2011.
[21] M. Rahman, S. Antani and G. Thoma, "A Learning-Based Similarity Fusion and Filtering Approach for Biomedical Image Retrieval Using SVM Classification and Relevance Feedback", *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, no. 4, pp. 640-646, 2011.
[22] M. Nazari and E. Fatemizadeh, "A CBIR System for Human Brain Magnetic Resonance Image Indexing", *International Journal of Computer Applications*, vol. 7, no. 14, pp. 33-37, 2010.
[23] W. Klecka, *Discriminant analysis*. Newbury Park, Calif.: Sage Publ., 2003.
[24] A. Smola and B. Schölkopf, "A tutorial on support vector regression", *Statistics and Computing*, vol. 14, no. 3, pp. 199-222, 2004.
[25] M. Kuhn and K. Johnson, *Applied predictive modeling*. New York: Springer, 2016.
[26] Z. Zhou, *Introduction to ensemble methods*. Boca Raton, Fla.: Chapman & Hall/CRC, 2012.
[27] D. Goldstein, "Analyzing Microarray Gene Expression Data", *Journal of the American Statistical Association*, vol. 100, no. 472, pp. 1464-1465, 2005.
[28] Corel 1000 image database [Online]. Available: http://wang.ist.psu.edu/docs/related/
[29] Vistex database [Online]. Available: http://vismod.media.mit.edu/pub/VisTex/
[30] The Database of Faces Cambridge University Computer Laboratory. [Online].Available: https://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html
[31] N. Fathi, B. Vaseghi and m. Shemshadi, "An overview of content-based image retrieval techniques (CBIR)", *IOSR Journal of Computer Engineering*, vol. 18, no. 04, pp. 108-116, 2016.